



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 99177

TO: Rebecca Prouty
Location: cm-1/10a13/10d01
Art Unit: 1652
Friday, August 01, 2003

Case Serial Number: 10/089986

From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534

toby.port@uspto.gov

Search Notes

Dear Examiner Prouty,

Here are the results of your search.

Please feel free to contact me if you have any questions.

Toby Port

BEST AVAILABLE COPY

Result No.	Score	Query Match	Length	DB ID	Description
1	74.4	8.1	1356	11	US-09-880-107-1648 Sequence 1648, App
2	44	4.8	421	11	US-09-960-353-6109 Sequence 6109, App
3	43.6	4.7	471	12	US-09-918-995-15653 Sequence 15653, A
4	39	4.2	377	11	US-09-960-352-8453 Sequence 8453, App
5	38.6	4.2	11812	15	US-10-239-676-209 Sequence 209, App
6	37.2	4.0	3290	15	US-10-027-632-116285 Sequence 116285, App
7	37.2	4.0	3290	15	US-10-027-632-116286 Sequence 116286, App
8	35.8	3.9	1114	15	US-10-099-278-85 Sequence 85, App1
9	35.8	3.9	1241	15	US-10-099-278-42 Sequence 42, App1
10	35.8	3.9	2132	15	US-10-099-278-39 Sequence 39, App1
11	35.2	3.8	32038	10	US-09-764-878-292 Sequence 292, App
12	35.2	3.8	32038	15	US-10-079-854-292 Sequence 292, App
13	35	3.8	520	14	US-10-146-731-144 Sequence 144, App
14	35	3.8	520	15	US-10-123-155-144 Sequence 144, App
15	34.8	3.8	630	11	US-09-874-300-2378 Sequence 2378, App
16	34.8	3.8	162450	15	US-10-071-179-1 Sequence 1, App1

QY 415 AGTTTATATCATTTAGGGGCTGATATAAAGTAAATATGACAGACACATAGCTATA 474
 Db 490 GCCCTGTACTCCTCTCTCTGATACACAGTGAAGAAAGTACTTTGACCAAGTGGACATT 549
 QY 475 GCAATAGGCTCGGCTGGAGTAATGTTGAAGAAATGTATTATATATTGTTATCGGGGAAA 534
 Db 550 TCCATAGTGGTGGATGGTGGCTAGACCAACAATCTTCTATTACATTTGACACCTGTCC 609
 QY 535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTTATCCATATCAATCAACAGGCCA 594
 Db 610 TACTCCGTGGATGCTTTGACTATGACCTGCAGACAGACAGATCTCCAAACCCGACAGAT 669
 QY 595 TTAATTTACTTTTGAAGACATGAATGCTTGGATATCCAGATGCTCAACAATGATGAG 654
 Db 670 GTTTACAAAGCTAGAAAAGGAAGAA-----CAAAATCCAGATGGAATGTATTGATGCT 723
 QY 655 GAGGTAATTTATGGTGGCTTTTCCAAAGCAGCGAATTTATTAATCACTACATCCAA 714
 Db 724 GAGGGAAGCTCTGGTGGCTGTTTACAATGGAGGAGTGAATCGTTTAGATCCGTG 783
 QY 715 CAACCGGAAGTGTACTGGATACCGTAAATACAGATCCTCAGTCACTCTGTAGCA 774
 Db 784 ACAGGAAA---AGACTTCAACTGTGAAGTGTGCTGTTGATAAAACAACCTTCAATGCTC 840
 QY 775 TTTGGCGTCCGAATTTGGATGAAGTGTGATGATGATGATGATGATGATGATGATGATG 834
 Db 841 TTTGGAGGAGAAATTAATCTGAAATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 835 AGTTCTTTNGACAAAGTTTATGTTAATGGGACGCTCTACAGATTAACAGGTTTGGCGTC 894
 Db 901 GAGGCTCTTTGAGCAACCTGAAGCTGGTGAATTTTCAAGATAAAGTGTCTGGGGTC 960
 QY 895 AAAGTTTTCG 905
 Db 961 AAAGGAATTCG 971

RESULT 2
 US-09-960-352-6109
 ; Sequence 6109, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Ningbing
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 ; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
 ; FILE REFERENCE: 16511.006/37-21(10298)C
 ; CURRENT APPLICATION NUMBER: US/09/960,352
 ; CURRENT FILING DATE: 2001-09-24
 ; NUMBER OF SEQ ID NOS: 15112
 ; SEQ ID NO 6109
 ; LENGTH: 421
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; OTHER INFORMATION: Clone ID: 26-LIB34-038-Q1-E1-G9
 US-09-960-352-6109

Query Match 4.8%; Score 44; DB 11; Length 421;
 Best Local Similarity 50.2%; Pred. No. 0.025;
 Matches 135; Conservative 0; Mismatches 131; Indels 3; Gaps 1;
 QY 631 CCAGATGGTCAACAATTCATGAGGAGGTAATTTATGGTTGCGGTTTTCACAGGACAG 690
 Db 116 CCAGATGGATGTGATTGATGTTGAGGGAAGCTCTGGTGGCTGTTTACATGGAGGA 175
 QY 691 CGAATTTATTAATACAGTACCCCAACACCGGAAGTGTACTGGATACCGTAAATACCA 750
 Db 176 AGAGTATCGGTTTGGATCCTGAGACAGGAAA---AGACTCCAAACTGTGAAGTTGCT 232
 QY 751 GATCCTCAGGTACCTCTGTAGCAATTTGGCGTCCGAATTTGGATCACTGATGAACA 810

Db 233 GTTGATAAAACAACCTCATCTGCTTCGGAGGGAAGGATTTACTCTGAAATGTACGTGACC 292
 QY 811 TCTGCTGGTCTTCAGCTTCACGACAGTCTTTTNGACAAAGTTTAGTTAATGGCACGTC 870
 Db 293 TGTGCCCGGGATGGGTGGATCCCAAGGGTCTTCTGCAACAACCTGAGGCTGGTGAATT 352
 QY 871 TACAGAGTAACAGGTTTAGCGCTCAAAAGG 899
 Db 353 TTCAAGATAACTGGCCTAAGAGTCAAAAG 381
 RESULT 3
 US-09-918-995-15653
 ; Sequence 15653, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 15653
 ; LENGTH: 471
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(471)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-918-995-15653

Query Match 4.7%; Score 43.6; DB 12; Length 471;
 Best Local Similarity 48.9%; Pred. No. 0.035;
 Matches 180; Conservative 0; Mismatches 179; Indels 9; Gaps 2;
 QY 434 CTGATAAAAGGTAATAAATGCACGAGACACATAGCTATAGCAAAATGGCTCGCGTGA 493
 Db 112 CTGATCACCAGCTGAAAAGTACTTTGACCAGGTGGACATTTCCAATGGTTTGGATTGGT 171
 QY 494 GTAATGATTTGAAGAAAATGTATTATATGTTGCGGGGAAAAGAGAGTAGACGAGTACG 553
 Db 172 CGGTAGACCAACAATCTTCTATTACATTTGACAGCCTGCTCTACTCCGTGGATGCCCTTG 231
 QY 554 ATTATGATGCTTCTACATTTCCATCAGCAATCAACGCCCATTTATTTTCAAAAGC 613
 Db 232 ACTATGACCTGCAGACAGGACAGATCTCCAAACCCGACAGTGTTTACAGCTAGANAAG 291
 QY 614 ATGAAGTGCCTGGATATCCAGATGGTCAAAACAATTTGATGAGGAGGGTAAATTTATGGGTTG 673
 Db 292 AAGAA-----CAAAATCCAGATGGAATGTATTGATGCTGAGGGGAAGCTCTGGGTGG 345
 QY 674 CCGTTTTCACAGGACAGCGAATTTATTAATACAGTACCCACACACCCGAACTGTTACTGG 733
 Db 346 CCGTTTTCACAGGACAGCGAATTTATTAATACAGTACCCACACACCCGAACTGTTACTGG 402
 QY 734 ATACCGTAAATAATACCAGATCCTCAGGTACCTCTGTAGCAATTTGGCGGTCCTGAAATTTGG 793
 Db 403 AACTGTGAATTTGCGCTGTTTATTAACAACACTTTCATGCTGCTTGGAGGGAAGATTTACT 462
 QY 794 ATGAATG 801
 Db 463 CTGAATG 470

RESULT 4
 US-09-960-352-8453
 ; Sequence 8453, Application US/09960352
 ; Patent No. US20020137139A1

QY 447 AAAATGCACGAGCAACATAGCTATAGCAAAATGGGCTCGCTGGAGTAATGATTTGAA 506
Db 705 AAAATATCAATATTTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCACATGAAA 764
QY 507 GAAATATGATTTATGATTCGGGGAAGAAAGAGTAGACAGTACGATTATGATGCTTC 566
Db 765 AAAATCATATCTTATGAATGGGCGCAGAAATAGATCAAAAACAACAATATAATATAC 824
QY 567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAGAAGCATGAAGTCCCTGG 626
Db 825 ATTAGGATTTCAAAATCAATATAGATTCAGGAATCAAATTTGATATACCAGAAGTAGTGG 884
QY 627 ATATCCAGATGGTCAACAATTTGATGAGGAGGTAATTTATGCGTTGCCGTTTCCAAAG 686
Db 885 AGGTACAGATGAATAAAAACAACAACATAATGAAGAATTTAAAATAGAATATAGATCATGA 944
QY 687 ACAGCGAATATTAAATTCAGTACCCCAACACCGGAAGT 725
Db 945 AACTAAATATGAAATAATATCAAGACAACTCTGAAAT 983

RESULT 10

US-10-099-278-39
; Sequence 39, Application US/10099278
; Publication No. US20030106093A1
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; APPLICANT: Michaels, Tracy E.
; APPLICANT: Finstad Lee, Stacey
; APPLICANT: Burmeister, Paula
; APPLICANT: Dojillo, Joanna
; TITLE OF INVENTION: Pesticidal Proteins
; FILE REFERENCE: MA-703C2D1
; CURRENT APPLICATION NUMBER: US/10/099,278
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 09/378,088
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: US 08/633,993
; PRIOR FILING DATE: 1996-04-19
; PRIOR APPLICATION NUMBER: US 08/844,188
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-099-278-39

Query Match 3.9%; Score 35.8; DB 15; Length 2132;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 127; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 447 AAAATGCACGAGCAACATAGCTATAGCAAAATGGGCTCGCTGGAGTAATGATTTGAA 506
Db 1401 AAAATATCAATATTTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCACATGAAA 1460
QY 507 GAAATATGATTTATGATTCGGGGAAGAAAGAGTAGACAGTACGATTATGATGCTTC 566
Db 1461 AAAATCATATCTTATGAATGGGCGACAGAAATAGATCAAAAACAACAATATAATATAC 1520
QY 567 TACATTATCCATCAGCAATCAACGCCATTTATTTACTTTTGAAGAAGCATGAAGTCCCTGG 626
Db 1521 ATTAGGATTTCAAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGTGG 1580
QY 627 ATATCCAGATGGTCAACAATTTGATGAGGAGGGTAATTTATGGGTTGCCGTTTCCAAAG 686
Db 1581 AGGTACAGATGAATAAANAACAACAACATAATGAAGAATTTAAAATAGAATATAGTATCA 1640

QY 687 ACAGCGAATTTAAATTCAGTACCCCAACACCGGAAGT 725
Db 1641 AACTAAATATGAAAAATATCAAGAACAATCTGAAAT 1679

RESULT 11

US-09-764-878-292
; Sequence 292, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
; LENGTH: 32038
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-292

Query Match 3.8%; Score 35.2; DB 10; Length 32038;
Best Local Similarity 49.5%; Pred. No. 84;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 473 TAGCAATGGCTCGCTGGAGTAATGATTTGAAGAAATGATTTATGATTCGGGGA 532
Db 17391 TAACTAATGACCATGGAGGGAGATTAATTTAAGAAAAATTAATCTTTTGTATTATTC 17450
QY 533 AAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGC 592
Db 17451 TAAAGGTAAAACTAAACCTTTAAGTGATTTAAAGTTTCATCATGAAGCAAGG 17510
QY 593 CATTTATTTCTTTTGAAGCATGAAGTGCCTGGATATCCAGATGGTCAACAAATTTGATG 652
Db 17511 CCAACAGCAGCTAAGAGATGTTTGAAGAATGAGGAAAGCTATTTGGCATACTAGATTA 17570
QY 553 AGGA 656
Db 17571 AAGA 17574

RESULT 12

US-10-079-854-292
; Sequence 292, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
; LENGTH: 32038
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-079-854-292

Query Match 3.8%; Score 35.2; DB 15; Length 32038;
Best Local Similarity 49.5%; Pred. No. 84;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 473 TAGCAATGGCTCGCTGGAGTAATGATTTGAAGAAATGATTTATGATTCGGGGA 532
Db 17391 TAACTAATGACCATGGAGGGAGATTAATTTAAGAAAAATTAATCTTTTGTATTATTC 17450

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:14:17 ; Search time 64 Seconds
(without alignments)
641.269 Million cell updates/sec

Title: US-10-089-986-2

Perfect score: 1615

Sequence: 1 MGPFVEKIALGKTYGEGP.....HVVRYTGLGVKGAGVKVL 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	308	22 AAB82087	Firefly luciferin
2	889.5	55.1	309	23 ABB09720	Amino acid sequenc
3	859.5	53.2	307	23 ABB06339	Luciola lateralis
4	568	35.2	303	22 ABB63233	Drosophila melanog
5	550	34.1	303	22 ABB58550	Drosophila melanog
6	550	34.1	303	22 ABB58551	Drosophila melanog
7	550	34.1	303	22 ABB66692	Drosophila melanog
8	457.5	28.3	299	23 ABB83787	Mouse SMP30 SEQ ID
9	454.5	28.1	299	16 AAR75416	Rat regucalcin, a
10	448.5	27.8	299	16 AAR74219	Human ageing marke

11	448.5	27.8	299	18 AAW14475	Human senility mar
12	448.5	27.8	299	20 AAY06995	Senescence marker
13	238	14.7	293	21 AAB11651	A. vitis hypersens
14	238	14.7	340	21 AAB11650	A. vitis hypersens
15	195.5	12.1	107	23 ABP32045	Human RNA polymera
16	124.5	7.7	320	23 ABP76347	Z. mobilis aldonoal
17	110	6.8	332	23 ABP39801	Staphylococcus epi
18	110	6.8	762	23 AAU78666	Human NOV5b protei
19	109	6.7	729	23 AAU78665	Human NOV5a protei
20	109	6.7	2009	22 ABB64069	Drosophila melanog
21	104	6.4	324	22 AAB99723	Staphylococcus aur
22	103.5	6.4	369	21 AAG07424	Arabidopsis thalia
23	103.5	6.4	371	21 AAG07423	Arabidopsis thalia
24	103.5	6.4	398	21 AAG07422	Arabidopsis thalia
25	97.5	6.0	553	21 AAY74581	Neisseria meningit
26	97.5	6.0	553	23 AAU72944	Neisseria meningit
27	97.5	6.0	695	23 AAU91286	Human NOV5e protei
28	97	6.0	371	21 AAG25423	Arabidopsis thalia
29	97	6.0	377	21 AAG25422	Arabidopsis thalia
30	96	5.9	419	22 ABB58148	Drosophila melanog
31	96	5.9	3117	21 AAY53667	Sequence gi/332818
32	95	5.9	265	22 AAG82348	S. epidermidis ope
33	94.5	5.9	553	21 AAY74580	Neisseria meningit
34	94	5.8	413	23 ABP40266	Staphylococcus epi
35	93	5.8	396	16 AAR76550	Thermotable alkali
36	93	5.8	449	23 ABP28423	Streptococcus poly
37	92.5	5.7	370	21 AAG06314	Arabidopsis thalia
38	92.5	5.7	385	21 AAG06313	Arabidopsis thalia
39	92	5.7	804	23 AAU91287	Human NOV5f protei
40	92	5.7	825	23 AAU91284	Human NOV5c protei
41	92	5.7	847	23 AAU91289	Human NOV5h protei
42	92	5.7	857	23 AAU91288	Human NOV5g protei
43	92	5.7	1012	23 AAU91285	Human NOV5d protei
44	92	5.7	4263	23 ABB97541	Novel human protei
45	92	5.7	4349	23 AAU79940	Human protocadheri

ALIGNMENTS

RESULT 1

AAB82087	ID	AAB82087 standard; Protein; 308 AA.
XX	AC	AAB82087;
XX	DT	26-JUN-2001 (first entry)
XX	DE	Firefly luciferin regenerating protein.
XX	KW	Firefly; luciferin regenerating protein; oxyluciferin; luciferase.
XX	OS	Photinus pyralis.
XX	FH	Key
XX	FT	Misc-difference 281
XX	XX	/note= "Encoded by TTN"
XX	PN	WO200125426-A1.
XX	PD	12-APR-2001.
XX	XX	22-SEP-2000; 2000WO-JP06527.
XX	PR	06-OCT-1999; 99JP-0285258.
XX	PA	(KIKK) KIKKOWAN CORP.
XX	PI	Kurosawa K, Kajiyama N;
XX	DR	WPI; 2001-266307/27.
XX	DR	N-PSDB; AAF86444.

PT Luciferin regenerating protein and gene encoding it useful for
 XX regenerating expensive luciferin from oxyluciferin and D-cysteine
 PS Claim 7; Page 16-17; 21pp; Japanese.
 CC The present sequence is a firefly protein which regenerates luciferin by
 CC using oxyluciferin and D-cysteine. This protein can be used for
 CC regeneration of luciferin, a substrate for luciferase, used for ATP
 CC (adenosine triphosphate) assays in both medical and food hygiene areas.
 XX Sequence 308 AA;
 SQ

Query Match 100.0%; Score 1615; DB 22; Length 308;
 Best Local Similarity 100.0%; Pred. No. 8.4e-153;
 Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPVVEKIAELGKTYVGEGRPHWDHETQTLVFDVTVETKTHKYPVPSQKKTFFCKVDKLVSF 60
 DB 1 MGPVVEKIAELGKTYVGEGRPHWDHETQTLVFDVTVETKTHKYPVPSQKKTFFCKVDKLVSF 60

QY 61 IIPLAGSPGRFVVSLEIREAILTWGVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNL 120
 DB 61 IIPLAGSPGRFVVSLEIREAILTWGVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNL 120

QY 121 WTGTMAIDAGLPVGTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSGKRR 180
 DB 121 WTGTMAIDAGLPVGTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSGKRR 180

QY 181 VDEYDYDASTLSISNQRPLFTFEKHEVPGPDGOTIDEENLWAVFQGORIHKISTQQP 240
 DB 181 VDEYDYDASTLSISNQRPLFTFEKHEVPGPDGOTIDEENLWAVFQGORIHKISTQQP 240

QY 241 EVLLDTVKIPDPQVTSVAFGPNLDLHVTSAGLQDDSSLDKSLVNGHVYRVTGLGVK 300
 DB 241 EVLLDTVKIPDPQVTSVAFGPNLDLHVTSAGLQDDSSLDKSLVNGHVYRVTGLGVK 300

QY 301 FAGVKVKL 308
 DB 301 FAGVKVKL 308

RESULT 2
 ABB09720
 ID ABB09720 standard; Protein: 309 AA.
 XX
 AC ABB09720;
 XX
 DT 11-JUN-2002 (first entry)
 XX
 DE Amino acid sequence of a protein capable of regenerating luciferin.
 XX
 DE Luciferin regenerating protein; luciferin; oxyluciferin; luminescence;
 KW ATP assay.
 XX
 OS Luciola cruciata.
 XX
 PN WO200210383-A1.
 XX
 PD 07-FEB-2002.
 XX
 PF 26-JUL-2001; 2001WO-JP06454.
 XX
 PR 28-JUL-2000; 2000JP-0228226.
 XX
 PA (KIKK) KIKKOMAN CORP.
 XX
 PI Hirokawa K, Kurosawa K, Kajiyama N;
 XX
 DR WPI: 2002-180080/23.
 DR N-PSDB; ABL41998.
 XX
 PT New Luciola cruciata-originated genes encoding proteins capable of
 PT regenerating luciferin especially from oxyluciferin, for producing

PT recombinant DNAs and transformants -
 XX Claim 1; Page 18-19; 23pp; Japanese.
 PS
 XX The present sequence represents a protein capable of regenerating
 CC luciferin. The protein can especially regenerate luciferin from
 CC oxyluciferin. The protein can be added to the luciferin-luciferase
 CC reaction system to sustain luminescence and reduce the amounts used.
 CC The polynucleotide sequence is useful for producing recombinant
 CC transformants, which are useful for the production of proteins useful
 CC in assaying ATP in medical sciences and food hygiene.
 XX Sequence 309 AA;
 SQ

Query Match 55.1%; Score 889.5; DB 23; Length 309;
 Best Local Similarity 56.6%; Pred. No. 2.8e-80;
 Matches 176; Conservative 47; Mismatches 83; Indels 5; Gaps 4;

QY 1 MGPVVEKIAELGKTYVGEGRPHWDHETQTLVFDVTVETKTHKYPVPSQKKTFFCKVDKLVSF 60
 DB 1 MGPVVEKIAELGKTYVGEGRPHWDHETQTLVFDVTVETKTHKYPVPSQKKTFFCKVDKLVSF 60

QY 61 IIPLAGSPGRFVVSLEIREAILTWGVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLG 118
 DB 61 IIPLAGSPGRFVVSLEIREAILTWGVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLG 120

QY 119 NLWTGTMAIDAGLPVGTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSG 177
 DB 121 NLWAGTNNGSDHTTGTTPVGTLSL-SNKQVKEHVSEVCISNGLAWSKDKKKFYVIDSA 179

QY 178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGPDGOTIDEENLWAVFQGORIHKIST 237
 DB 180 VRQVDQDFEDAKNLSNRQPLFTFEKHEVPGPDGOTIDEENLWAVFQGORIHKIST 239

QY 238 QPPEVLLDTVKIPDPQVTSVAFGPNLDLHVTSAGLQDDSSLDKSLVNGHVYRVTGLG 297
 DB 240 STPETLLGIVEIPEHQVTSVCIGGAELNVLIVTASIKLPGADETKPM-KGAIYKVTGLG 298

QY 298 VKGFAGVKVKL 308
 DB 299 VKGLPGDRVKL 309

RESULT 3
 ABB06339
 ID ABB06339 standard; Protein: 307 AA.
 XX
 AC ABB06339;
 XX
 DT 29-MAY-2002 (first entry)
 XX
 DE Luciola lateralis luciferin regenerating protein SEQ ID NO:2.
 XX
 DE Luciola lateralis; luciferin; regeneration; Japanese firefly;
 KW adenosine triphosphate; ATP; medical science; food hygiene;
 KW luciferase; luminescence.
 XX
 OS Luciola lateralis.
 XX
 PN WO200210384-A1.
 XX
 PD 07-FEB-2002.
 XX
 PF 26-JUL-2001; 2001WO-JP06455.
 XX
 PR 28-JUL-2000; 2000JP-0228227.
 XX
 PA (KIKK) KIKKOMAN CORP.
 XX
 PI Hirokawa K, Kurosawa K, Kajiyama N;
 XX
 DR WPI: 2002-227086/28.
 DR N-PSDB; ABL49712.


```

XX  Luciola lateralis-originated genes encoding proteins capable of
PT  regenerating luciferin especially from oxyluciferin, useful for
PT  producing recombinant DNAs and transformants to give proteins useful in
PT  assaying adenosine triphosphate
XX
XX  Claim 1; Page 18-19; 23pp; Japanese.
PS
XX  The present sequence represents a protein capable of regenerating
CC  luciferin, which is isolated from Luciola lateralis (Japanese firefly).
CC  The gene encoding the protein capable of regenerating luciferin can be
CC  used for producing recombinant DNAs and transformants, which can be
CC  used for the production of proteins useful in assaying adenosine
CC  triphosphate (ATP) in medical sciences and food hygiene. The protein
CC  can be added to the luciferin-luciferase reaction system to sustain
CC  luminescence and reduce their amounts used.
XX
XX  Sequence 307 AA;
SQ
Query Match 53.2%; Score 859.5; DB 23; Length 307;
Best Local Similarity 52.8%; Pred. No. 2.7e-77;
Matches 163; Conservative 58; Mismatches 85; Indels 3; Gaps 3;
QY 1 MGPVVEKTAELGKVTGEGPHWDHETQTLFYVDVTEKTFHKYVPSQKKYTFCKVDKLYSF 60
DB 1 MSPVIEQTEVDNFQIGEGPHWDHETQSLYFVDILEKSIHKYVPSTKQHTKMLNKRPSF 60
QY 61 IIPLAGSPGRVWSLEERIEAILTWGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120
DB 61 IIPKEYSRDRVISLERIDICVLTWGVSATSHELTIVDTGIEGTFNDGKADAFGNL 120
QY 121 WTGTMAIDAGL-PVGPVTGSLYHLAGDKKKVHMESNIAIANGLAWSNDLKKMYIDSQKR 179
DB 121 WAGTLYSKFDIEKQGNPTGLYSL-SNKQLRKHISNIFLSNGLAWNKKSKKFFYFDSNKR 179
QY 180 RVDEVDYDASTLSISNORPLTFEKEHVPYDQGTIDEENLWVAVFQGORIKISTQQ 239
DB 180 TIDQFDYDSENLIIISNCPLETFDQKHGQGLPDAQTIDENLWVAIVRGKVINIGTKQ 239
QY 240 PEVLDTYVKIPDPQVTSVAFGPNLDLHVTSAGLQDSSLDKSLVNGHVRVYRTGLGVK 299
DB 240 PESLGLVINMPESLITSCVFGSKLDLYVTTSIGKEYETDSTK-LYKGGLYRVTGLGVK 298
QY 300 GFAGVKVKL 308
DB 299 GLPAHREFSL 307
RESULT 4
ABB63233
ID ABB63233 standard; Protein; 303 AA.
XX
XX ABB63233;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 16491.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
FN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.

```

```

XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL07336.
DR
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 16491; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 303 AA;
SQ
Query Match 35.2%; Score 568; DB 22; Length 303;
Best Local Similarity 40.7%; Pred. No. 3.7e-48;
Matches 121; Conservative 48; Mismatches 116; Indels 12; Gaps 5;
QY 16 VEGPHWDHETQTLFYVDVTEKTFHKYVPSQKKYTFCKVDK--LVSFIIPLAGSPGRFVW 73
DB 15 LGEQPHWDVDRSLYVVDLESAGINRYDFKQKVVRAKIEGIEFASFILPVENKQPEFV 74
QY 74 SLEREIALTWGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDAGLPV 133
DB 75 CGGLRTVIVQDGVSAVAKVTRTTFEQVQDLKENRLNDAKTDPNGRFYGGTMA-DSSGDI 133
QY 134 GPVTSGLYHLAGDKKKVHMESNIAIANGLAWSNDLKKMYIDSQKKRRVDEYDYDASTLSI 193
DB 134 TQWKGELYSWQAGGPNNAIRSKVGLISNGLANDVRAKKFFYFIDTNNHEVLAYDYNOSTGAV 193
QY 194 SNQRLFTFEK--HEVPGYDQGTIDEENLWVAVFQGORIKISTQOPEVLLDTFVKIPD 251
DB 194 SNPKVIFDLRKIRPEGLPFGMTVDGNIYVATFNGGTVFKVNPSTGKILLE-IKIP 252
QY 252 PQVTSVAFGPNLDLHVTSAGLQDSSLDKSLVNGHVRVYRTGLGVKGFAGVKVKL 308
DB 253 TQITSVAFGPNLDLIVYTTA-----NKFDQPKPAGTTFQVTLNKAAGVNLKI 303
RESULT 5
ABB58550
ID ABB58550 standard; Protein; 303 AA.
XX
XX ABB58550;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 2442.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
FN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX (PEKE ) PE CORP NY.

```

```
PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL02653.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 2442; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 303 AA;
XX
XX Query Match 34.1%; Score 550; DB 22; Length 303;
XX Best Local Similarity 39.2%; Pred. No. 2.3e-46;
XX Matches 118; Conservative 48; Mismatches 115; Indels 20; Gaps 6;
QY 16 VEGPHDHEOTLTLYFVDTVEKTHKYPVSKKVTFCV--DKLVSFIIPLAGSPGRFV 73
DB 15 LGEAPHWDVARQSLYYVDLEAGSLRLRYDAQNKVKYTKIEGETLAGFVLPVEGRPQEFV 74
QY 74 SLERETAILTWDGVSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLWMTGTMAIDAGLPV 133
DB 75 GCGRRVIVNDGVSFSAKVVRTLEFVQPLMEKNRLNDAKVDPGRFPGGTMRV-----I 129
QY 134 GPV-----TGSYLHLGADKKVKMHESNIAIANGLAWSNDLKKMYIIDSKRRVDEYDAS 189
DB 130 GDEFEFRHGLYRWEAGGVSVIKGDVGISNGLAWDEKAKKFFYYIDTDDYEVKSYDYDFE 189
QY 190 TGISNORPLFTFEKHEVPGY--PDGOTIDEGNWLWAVFQGORIIKISTQOPEVLLDTV 247
DB 190 TGVASNPVKYIFNLKNSPKDHLDPGLTIDTEGNLYVATENGATIIYKVNPTGKILLE-I 248
QY 248 KIPDPQVTSVAFGGPNLDELHVTSAQLQDDSSLDKSLVNGHVYRVTGLGVKGAGVKVK 307
DB 249 KFTPTQITSAAGFGPNLDILYVTTA-----AKFDQAPAGTYYKVTGLNATGYPGVNLK 302
XX
XX RESULT 6
XX ID ABB58551 standard; Protein; 303 AA.
XX AC ABB58551;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 2445.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
```

```
PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL02654.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 2445; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 303 AA;
XX
XX Query Match 34.1%; Score 550; DB 22; Length 303;
XX Best Local Similarity 39.2%; Pred. No. 2.3e-46;
XX Matches 118; Conservative 48; Mismatches 115; Indels 20; Gaps 6;
QY 16 VEGPHDHEOTLTLYFVDTVEKTHKYPVSKKVTFCV--DKLVSFIIPLAGSPGRFV 73
DB 15 LGEAPHWDVARQSLYYVDLEAGSLRLRYDAQNKVKYTKIEGETLAGFVLPVEGRPQEFV 74
QY 74 SLERETAILTWDGVSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLWMTGTMAIDAGLPV 133
DB 75 GCGRRVIVNDGVSFSAKVVRTLEFVQPLMEKNRLNDAKVDPGRFPGGTMRV-----I 129
QY 134 GPV-----TGSYLHLGADKKVKMHESNIAIANGLAWSNDLKKMYIIDSKRRVDEYDAS 189
DB 130 GDEFEFRHGLYRWEAGGVSVIKGDVGISNGLAWDEKAKKFFYYIDTDDYEVKSYDYDFE 189
QY 190 TGISNORPLFTFEKHEVPGY--PDGOTIDEGNWLWAVFQGORIIKISTQOPEVLLDTV 247
DB 190 TGVASNPVKYIFNLKNSPKDHLDPGLTIDTEGNLYVATENGATIIYKVNPTGKILLE-I 248
QY 248 KIPDPQVTSVAFGGPNLDELHVTSAQLQDDSSLDKSLVNGHVYRVTGLGVKGAGVKVK 307
DB 249 KFTPTQITSAAGFGPNLDILYVTTA-----AKFDQAPAGTYYKVTGLNATGYPGVNLK 302
XX
XX RESULT 7
XX ID ABB66692 standard; Protein; 303 AA.
XX AC ABB66692;
XX
XX DT 26-MAR-2002 (first entry)
XX
```


ID AAR75416 standard; Protein; 299 AA.
 AC AAR75416;
 XX
 DT 31-JAN-1996 (first entry)
 DE Rat regucalcin, a calcium binding protein.
 XX
 KW Calcium binding protein; regucalcin; rat.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 95
 FT /note= "corresponds to CAA codon"
 FT Misc-difference 141
 FT /note= "corresponds to GAC codon"
 FT
 PN JP07123985-A.
 PD 16-MAY-1995.
 XX
 XX 09-NOV-1993; 93JP-0279349.
 XX
 XX 09-NOV-1993; 93JP-0279349.
 XX
 PA (DAI-) DAICHI KAKAGU YAKUHI N K.
 PA (YAMA/) YAMAGUCHI M.
 XX
 DR WPI: 1995-211628/28.
 DR N-PSDB; AAR7295.
 XX
 PT DNA encoding regucalcin - useful for the recombinant production of
 PT regucalcin, a calcium binding protein
 XX
 PS Claim 1; Page 5-6; 6pp; Japanese.
 XX
 CC A rat liver cDNA clone coding for a calcium binding protein of mol.
 CC wt. 33388 was isolated and sequenced (AAR7295). The protein was
 CC designated "regucalcin" (AAR75416).
 XX
 SQ Sequence 299 AA;
 Query Match 28.1%; Score 454.5; DB 16; Length 299;
 Best Local Similarity 37.5%; Pred. No. 8e-37;
 Matches 111; Conservative 41; Mismatches 121; Indels 23; Gaps 7;
 QY 14 YTVGEGPHDHTOTLYFVDVTEKTHFYVPSQKYYTECKVDKLVSLFIPLAGSPGRV 73
 DB 14 YRCGSPVWEAEASKCLLFIDIPSKTVCRWDSISNRVQGVDPVSSV--ALROSGGYA 71
 QY 74 SLERETAILTWDCVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDAGLPV 133
 DB 72 TIGTKFCALNWE-----DQSFILAMVEDKNNRNDGKVDPAGRYFAGTMAEE---T 122
 QY 134 GPVT-----GSLYHLGADKKVHMESNTAIAIANGLAWSNDLKKMYIDSGKRRVDEYDVA 188
 DB 123 APAVLERHQSLSLFPDHSVKKYFNQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182
 QY 189 STLSTISNORPLTFEKEHVEPGYPDGOTIDEENLWAVFOGQRIIKISTQOPEV--LLDT 246
 DB 183 PTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKLWVACNGRVRRL--DPETGKRLQT 237
 QY 247 VKIPDPQVTSVAFGGPNLDLHVTSAQLQDSSLDKSLVNGHVVYRVTCGLGVKGA 302
 DB 238 VKLPVDKTTSCCGGKGYSEMYVTCTARDGMSAEGLLRQPDAGNIFKITGLGVKGA 293
 RESULT 10
 AAR74219
 ID AAR74219 standard; Protein; 299 AA.
 XX
 AC AAR74219;

XX 03-JAN-1996 (first entry)
 DT
 XX Human ageing marker protein, SMP30.
 DE
 XX SMP30; marker protein; ageing; organ development.
 KW
 XX Homo sapiens.
 OS
 XX JP07097399-A.
 PN
 XX 11-APR-1995.
 PD
 XX 29-SEP-1993; 93JP-0265681.
 XX
 PF 29-SEP-1993; 93JP-0265681.
 XX
 PR 29-SEP-1993; 93JP-0265681.
 XX
 XX (EJRE) FUJI REBIO KK.
 PA
 XX WPI: 1995-175363/23.
 XX N-PSDB; AAR90035.
 DR
 DR Novel polypeptide for detecting human ageing marker protein SMP30 -
 PT for monitoring liver and kidney development in new-born babies
 PT
 XX Claim 1; Page 7-8; 10pp; Japanese.
 PS
 XX AAQ90035 is the SMP30 gene. It encodes the human ageing marker protein,
 CC SMP30 (AAR74219). Human SMP30 is found in human organs, tissues, blood,
 CC urine and cerebrospinal fluid. The blood concentration of SMP30 is
 CC known to increase with renal and hepatic deficiencies and to decrease
 CC with age. It is therefore useful in the monitoring of renal or hepatic
 CC deficiencies and for the monitoring of the development of the liver
 CC and kidneys in newborn babies.
 XX
 SQ Sequence 299 AA;
 Query Match 27.8%; Score 448.5; DB 16; Length 299;
 Best Local Similarity 36.9%; Pred. No. 3.2e-36;
 Matches 110; Conservative 42; Mismatches 113; Indels 33; Gaps 8;
 QY 17 GEGPHDHTOTLYFVDVTEKTHFYVPSQKYY---TFCKVDKLVSLFIPLAG---SPG 69
 DB 17 GEGPVWEVSNLSLFDV-----IPAKKVCRWDSFTKQVORVTMDAPVSSVALRQSG 67
 QY 70 RFVVSLEREITAILTWDCVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDA 129
 DB 68 GYVATIGTKFCALNWEKQSAV-----VLATVDNDKNNRNDGKVDPAGRYFAGTMAEE- 121
 QY 130 GLPVGPPVT-----GSLYHLGADKKVHMESNTAIAIANGLAWSNDLKKMYIDSGKRRVDEY 184
 DB 122 ---TAPAVLERHQSLSLFPDHSVKKYFNQVDISNGLDWSLDHKIFYYIDSLSYSDAF 178
 QY 185 DYDASTLSISNORPLTFEKEHVEPGYPDGOTIDEENLWAVFOGQRIIKISTQOPEVLL 244
 DB 179 DYDLOQTQISNRRTVYKLEKEE--QIPDGMCIDAEGLWVACNGRVRILDPVTGK-RL 235
 QY 245 DTVKIPDPQVTSVAFGGPNLDLHVTSAQLQDSSLDKSLVNGHVVYRVTCGLGVKGA 302
 DB 236 QTVKLPVDKTTSCCGGKGYSEMYVTCTARDGMDPEGLLROPEAGGIFKITGLGVKGA 293
 RESULT 11
 AAR14475
 ID AAR14475 standard; peptide; 299 AA.
 XX
 AC AAR14475;
 XX
 DT 28-MAY-1997 (first entry)
 DE Human senility marker protein, hSMP30.
 XX
 KW SMP30; senility marker protein; monoclonal antibody; detection.

```

XX OS Homo sapiens.
XX PN JP08319298-A.
XX PD 03-DEC-1996.
XX PF 25-MAY-1995; 95JP-0149791.
XX PR 25-MAY-1995; 95JP-0149791.
XX PA (FJRE ) FUJI REBIO KK.
XX DR WPI; 1997-073109/07.
XX PT Anti-human senility marker protein monoclonal antibody - useful for
XX detection of protein
XX PS Claim 3; Page 6-7; 8pp; Japanese.
XX CC The present sequence is that of human senility marker protein (hSMP30)
XX which has a molecular weight of 30 kDa. Monoclonal antibodies recognising
XX hSMP30 are claimed and can be used in a method for detection of the
XX hSMP30 protein in a sample.
XX SQ Sequence 299 AA;
    Query Match 27.8%; Score 448.5; DB 18; Length 299;
    Best Local Similarity 36.9%; Pred. No. 3.2e-36;
    Matches 110; Conservative 42; Mismatches 113; Indels 33; Gaps 8;
Qy 17 GEGPHDHTQTLFYVDVTEKTHKYPVPSOKY---TFCKVDKLVFIPLAG----SPG 69
Db 17 GESPWEEVSNLLFVD-----IPAKKVCRWDSFTKQVQRTMDAPVSSVALRQSG 67
Qy 70 RFVVSLERETAILTWDGVSAAPTSIETIAIVNVEPHIKNNRLNDGKADPLGNLTGTMADA 129
Db 68 GYVATIGTKFCALNWKQSAV-----VLATVDNDKNNRFDGKVDPAGRYFAGTMAEE- 121
Qy 130 GLPVGPIVT-----GSLYHLGADKKVKMHESNIAIANGLANSLDKMKWYIDSGKRRVDEY 184
Db 122 ---TAPAVLERHQGALYSLPDHHVKYFDQVDISNGLDWSLDHKKIFYIDSLSYSVDAF 178
Qy 185 DYDASTLSISNORPLTFTEKHEVPGYDGTIDEENLWAVFQGGRIKISTOOPEVLL 244
Db 179 DYDLQTGOISNRRSVYKLEKEE--QIPDGMCIDAEGLWVACYNNGRIVRLDPVTGK-RL 235
Qy 245 DTVKIPDPQVTSVAFGGPNLDLHVTAGLQDSSLDKSLVNGHVYRVTGLGVKGEA 302
Db 236 QTVKLPVDKTTSCCFGGKNYSYMTVCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293

RESULT 12
AAY06995
ID AAY06995 standard; Protein; 299 AA.
XX AC AAY06995;
XX DT 02-JUL-1999 (first entry)
XX DE Senescence marker protein SMP30.
XX KW Cancer associated antigen; diagnosis; research; treatment; human;
XX breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
XX prostate cancer.
XX OS Homo sapiens.
XX PN WO9904265-A2.
XX PD 28-JAN-1999.
XX PF 15-JUL-1998; 98WO-US14679.

```

```

XX 22-JUN-1998; 98US-0102322.
XX PR 17-JUL-1997; 97US-0896164.
XX PR 10-OCT-1997; 97US-0061599.
XX PR 10-OCT-1997; 97US-0061765.
XX PR 10-OCT-1997; 97US-0948705.
XX PR 11-OCT-1997; 97GB-0021697.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX XX
XX Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
XX Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
XX Tureci O;
XX WPI; 1999-132448/11.
XX N-PSDB; AAX40196.
XX
XX New isolated cancer associated nucleic acids and polypeptides -
XX for the diagnosis, monitoring or treatment of cancers
XX
XX Example 8; Page 775-776; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
XX by expression of a human cancer associated antigen precursor coded for by
XX a nucleic acid molecule (NAM). The method comprises: (a) contacting a
XX biological sample isolated from a subject with an agent that specifically
XX binds to the NAM, an expression product or a fragment of an expression
XX product complexed with an HLA molecule; and (b) determining the
XX interaction between the agent and the NAM or the expression product as a
XX determination of the disorder. The products and methods can be used in
XX the diagnosis, monitoring, research, or treatment of conditions
XX characterised by the expression of various cancer associated antigens.
XX The invention provides nucleic acid sequences and encoded polypeptides
XX which are cancer associated antigen precursors expressed in human breast
XX cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
XX lung cancer.
XX SQ Sequence 299 AA;
    Query Match 27.8%; Score 448.5; DB 20; Length 299;
    Best Local Similarity 36.9%; Pred. No. 3.2e-36;
    Matches 110; Conservative 42; Mismatches 113; Indels 33; Gaps 8;
Qy 17 GEGPHDHTQTLFYVDVTEKTHKYPVPSOKY---TFCKVDKLVFIPLAG----SPG 69
Db 17 GESPWEEVSNLLFVD-----IPAKKVCRWDSFTKQVQRTMDAPVSSVALRQSG 67
Qy 70 RFVVSLERETAILTWDGVSAAPTSIETIAIVNVEPHIKNNRLNDGKADPLGNLTGTMADA 129
Db 68 GYVATIGTKFCALNWKQSAV-----VLATVDNDKNNRFDGKVDPAGRYFAGTMAEE- 121
Qy 130 GLPVGPIVT-----GSLYHLGADKKVKMHESNIAIANGLANSLDKMKWYIDSGKRRVDEY 184
Db 122 ---TAPAVLERHQGALYSLPDHHVKYFDQVDISNGLDWSLDHKKIFYIDSLSYSVDAF 178
Qy 185 DYDASTLSISNORPLTFTEKHEVPGYDGTIDEENLWAVFQGGRIKISTOOPEVLL 244
Db 179 DYDLQTGOISNRRSVYKLEKEE--QIPDGMCIDAEGLWVACYNNGRIVRLDPVTGK-RL 235
Qy 245 DTVKIPDPQVTSVAFGGPNLDLHVTAGLQDSSLDKSLVNGHVYRVTGLGVKGEA 302
Db 236 QTVKLPVDKTTSCCFGGKNYSYMTVCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293

RESULT 13
AAB11651
ID AAB11651 standard; Protein; 293 AA.
XX AC AAB11651;
XX DT 23-OCT-2000 (first entry)
XX

```



```

XX SQ Sequence 340 AA;
Query Match 14.7%; Score 238; DB 21; Length 340;
Best Local Similarity 26.5%; Pred. No. 4.2e-15;
Matches 80; Conservative 46; Mismatches 126; Indels 50; Gaps 12;

Qy 16 VGEQPHWDHETQTYFYVDVTEKTF--HKYVPSQKKYTFCKVDKLVFIPIAGSPGRFV 73
Db 64 LGEQTYDVARDTAWFEDILGKLIHRFATGE-----TIRHDLPMNAS-ALATI 112
Qy 74 SLREIETAILTWGCV--SAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLTGTMADAG 130
Db 113 DAERQM-LATEQGIFIRTIASGETLTLTALEPERPGRNSNDGRVHPGCAIMGTM---G 167
Qy 131 LPVGPVTSGLYHLGADKKVKMHESNLAIAANGLAWSNDLKKNMYIDSCKRRVDEYDYDAST 190
Db 168 KTAADGAGAIYHVAGDKVTRLYDS-ISPNAICFSPDGLSGLYYDTRVNTLMKSLDPAT 226
Qy 191 LSIISNQRPLTFEHEVPGYDQGTIDEENLWAVF-----QCQRILKISTQOP 240
Db 227 -GLPTGPTVHIDGRGKGGIDGAVCAEGGLNARWGVAVDHYDRTGQHLAR----- 279
Qy 241 EVLLDTVKIPDPQVTSVAFSGPNDELHVTSA--GLQLDSSLDKSLVNGHVYRVTLGV 298
Db 280 -----YELPAAQTCTAFIAGAKADRLLVTSATGLDAGLAADP---HGKTLVLDITV 330
Qy 299 KG 300
Db 331 KG 332

RESULT 15
ABP32045
ID ABP32045 standard; Protein: 107 AA.
AC ABP32045;
XX 09-JUL-2002 (first entry)
XX Human RNA polymerase-like ORF1018 protein, SEQ ID NO:2036.
XX Human: ORF: open reading frame; ORFX: drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX Homo sapiens.
XX WO200190366-A2.
XX 29-NOV-2001.
XX 24-MAY-2001; 2001WO-US17076.
XX 24-MAY-2000; 2000US-206690P.
XX (CURA-) CURAGEN CORP.
XX Leach MD, Shimkets RA;
XX WPI; 2002-106200/14.
XX DR N-PSDB; ABN76071.
XX

```

```

PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation
XX Claim 10; Page 773; 2508pp; English.
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN75987 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,
XX cell differentiation, immune modulation, haematopoiesis regulation,
XX tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
XX chemokinetic activity, haemostatic activity, thrombolytic activity,
XX receptor/ligand, antiinflammatory activity, tumour inhibition activity,
XX and antiinfective activity, and may also be involved in the determination
XX of bodily characteristics, fertility and behaviour. ORFX proteins,
XX nucleic acids and antibodies may be used in the treatment of cancers,
XX other proliferative disorders such as psoriasis and benign tumours,
XX neurological disorders such as epilepsy and Alzheimer's disease,
XX cardiovascular diseases, immune system disorders, disorders related to
XX organ transplantation, disorders of tissue growth and regeneration,
XX diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
XX storage disease, and infectious diseases caused by viral, bacterial,
XX fungal and other pathogens. ORFX nucleic acids may also be used as a
XX source of primers and probes, in the detection of ORFX genomic sequences
XX or transcripts, in the identification and cloning of homologous
XX sequences, in genetic diagnosis, and in forensic biology. The ORFX
XX nucleic acids may additionally be used to produce transgenic animals
XX which may be useful for studying the function and/or activity of ORFX
XX protein, and in drug screening. The ORFX proteins may also be used as
XX immunogens to generate specific antibodies, which are useful in the
XX diagnosis, treatment and monitoring of ORFX-associated diseases.
XX Sequence 107 AA;
Query Match 12.1%; Score 195.5; DB 23; Length 107;
Best Local Similarity 43.4%; Pred. No. 1.3e-11;
Matches 49; Conservative 9; Mismatches 44; Indels 11; Gaps 3;
Qy 115 DPLGNLTGTMADAGLPVGPVT-----GSLYHLGADKKVKMHESNLAIAANGLAWSNDLK 169
Db 1 DPAGRYTAGTMAEE-----TAPAVLERHQGSLYSLFPDHSVKRYFDQYDINSGLDWSLDHK 56
Qy 170 KMYYIDSGKRVRDEYDYDASTLSISNORPLTFEKEHVEVPGYDQGTIDEENL 222
Db 57 IFYYIDSLSYTVXAFDVLDTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKL 107
Search completed: August 1, 2003, 14:32:07
Job time : 67 secs

```


Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	550	34.1	303	2	JC7250	anterior fat body
2	538	33.3	306	2	JC7249	anterior fat body
3	457.5	28.3	299	2	S72173	sensence marker
4	454.5	28.1	299	2	S34588	sensence marker
5	448.5	27.8	299	1	S60035	sensence marker
6	423.5	26.2	569	2	E95309	probable calcium b
7	409.5	25.4	560	2	A13216	hypothetical prote
8	376	23.3	292	1	A70047	RNA polymerase hom
9	365.5	22.6	285	2	B90445	regucalcin homolog
10	331	20.5	289	2	A87351	Smp-30/Cgr1 family
11	315	19.5	293	2	A87473	Smp-30/Cgr1 family
12	292	18.1	285	2	G90485	conserved hypothet
13	288.5	17.9	311	2	D95265	probable regucalc
14	273	16.9	293	2	AB3071	calcium-binding pr
15	273	16.9	322	2	G98215	sensence marker
16	256	15.9	295	2	AH2662	calcium-binding pr
17	256	15.9	314	2	G97444	probable calcium-b
18	206	12.8	294	2	AF3043	calcium-binding pr
19	206	12.8	305	2	F98242	regucalcin (AB0379
20	203	12.6	302	2	AH0832	conserved hypothet
21	202.5	12.5	291	2	AG3501	gluconolactonase (
22	153.5	9.5	306	2	AG3051	gluconolactonase p
23	153.5	9.5	348	2	D98234	probable gluconola
24	140	8.7	303	2	C95896	probable gluconola
25	132	8.2	304	2	G95274	probable gluconola
26	124.5	7.7	320	2	S28218	gluconolactonase (
27	120	7.4	347	2	B87401	gluconolactonase (
28	113	7.0	303	2	G87512	gluconolactonase,
29	111	6.9	356	2	D83643	conserved hypothet

C:Species: Sarcophaga peregrina
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: JG7249; PC7071
R:Nakajima, Y.; Natori, S.
J. Biochem. 127, 901-908, 2000
A:Title: Identification and characterization of an anterior fat body protein in an insect
A:Reference number: JG7249
A:Accession: JG7249
A:Molecule type: mRNA
A:Residues: 1-306 <NAK>
A:Cross-references: DDBJ:AB036903
A:Experimental source: larval fat body
A:Accession: PC7071
A:Molecule type: protein
A:Residues: 157-168;174-185;186-195 <NA2>
C:Comment: This protein, a homolog of the senescence marker protein-30, has its function in senescence.
C:Genetics:
A:Gene: afp
C:Superfamily: senescence marker protein-30

Query Match 33.3%; Score 538; DB 2; Length 306;
Best Local Similarity 39.0%; Pred. No. 6.1e-36;
Matches 122; Conservative 48; Mismatches 123; Indels 20; Gaps 7;

QY 3 PVVEKIAELGKVTYGEHPHMDHETQTLFYFVDVTEKTFHKYVPSQKKTFFCKV--DKLVSF 60
DB 7 PLPDSHAEL-----GEGPHMDVATONLYYVDINAGKLLRYNENKYYKAKIEGEDLAGF 61
QY 61 IPLAGSGPRFVSLEREIALTWDGVSAAPTSTEAIVNV---EPHKNRNLNDGKADPL 117
DB 62 ALPVEGTTDQFAVSGRRVIVQWDGISTAKVTKTFLFEVQTGDERFNGRFDNGKCDPR 121
QY 118 GNLWTGTAIDAGLPVGPVTSGLYHLGADKKVKHESNIAIANGLAWSNDLKKMYIDSG 177
DB 122 GRLEFAGTMY-VGDEFEHRYGELYKEGGKIEVVKSVNGVSNGLAWNEKTKKYYIDTT 180
QY 178 KRRVDEYDYDASTLSISNORPLTFEKEHVPFY--PDGQTIDEENLWAVFQGRITKI 235
DB 181 DYEVEKYDYDFETGATNPKVNVENLRKNSPKDHLPLDGMTIDTDCGNIVVATFNGYTIKV 240
QY 236 STQQPEVLLDTVKIPDPQVTSVAFGPNLDELHVTISAGLDLDDSSLDKSLYNGHVYRTG 295
DB 241 NPTTKGVLLLE-TKFFCKQITSAAGGPNLDILYVTTSS-----SRFEGPDPAAGTYYKVTG 293
QY 296 LGVKGFGAGVKVL 308
DB 294 LGAKGYPTWKIQL 306

RESULT 3
S72173
senescence marker protein 30 - mouse
N:Alternate names: regucalcin
C:Species: Mus musculus (house mouse)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Sep-1999
C:Accession: S72173; S72174
R:Fujita, T.; Shirasawa, T.; Maruyama, N.
Biochim. Biophys. Acta 1308, 49-57, 1996
A:Title: Isolation and characterization of genomic and cDNA clones encoding mouse senescence marker protein-30
A:Reference number: S72173; MUID:96328264; PMID:8765750
A:Accession: S72173
A:Molecule type: mRNA
A:Residues: 1-299 <FUJ>
A:Cross-references: EMBL:U28937; NID:g1143999; PIDN:AAC52721.1; PID:g1144000
A:Accession: S72174
A:Molecule type: DNA
A:Residues: 1-54 <FUW>
A:Cross-references: EMBL:U32170; NID:g1144337; PIDN:AAD03478.1; PID:g1144338
A:Experimental source: liver
C:Function:
A:Description: calcium binding
A:Note: age-associated decrease
C:Superfamily: senescence marker protein-30

C:Keywords: calcium binding

Query Match 28.3%; Score 457.5; DB 2; Length 299;
Best Local Similarity 37.8%; Pred. No. 1.9e-29;
Matches 112; Conservative 40; Mismatches 121; Indels 23; Gaps 7;

QY 14 YTVGEGPHMDHETQTLFYFVDVTEKTFHKYVPSQKKTFFCKV--DKLVSF 73
DB 14 YRCGESPVWEEASQSLLFVDIPSKIICRWDTVSNQVRVADAPVSSV--ALRQLGGYVA 71
QY 74 SLEREIALTWDGVSAAPTSTEAIVNVPEPHIKNNRNLNDGKADPLGNLWGTGTMAIDAGLPV 133
DB 72 TIGTKFCALNWE-----QSVFLAMVDEDDKNNRFDNKGKVDPAGRYFAGTMAEE---T 122
QY 134 GPVT-----GSLYHLGADKKVKHESNIAIANGLAWSNDLKKMYIDSGRRRVDEYDYDA 188
DB 123 APAVLERHQGLSYLSLPDHSVKKYFDQVDSISNGLDWSLDHKIFYYIDLSYTVDAFDYDL 182
QY 189 STLSISNORPLTFEKEHVPFYPDGQTIDEENLWAVFQGRITKIISTQOPEV--LLDT 246
DB 183 QTGQISNRRIVYKMEKDE--OIPDGMCIDAEGLWVACYNNGRVRIL---DPETGKRLOT 237
QY 247 VKIPDPQVTSVAFGPNLDELHVTISAGLDLDDSSLDKSLYNGHVYRTGLGVKGEA 302
DB 238 VKLPVDKTTSCFCGKDYSEMYVTCCARDGLNAGELLRQPDAGNIFKITGLGVKGA 293

RESULT 4
S34588
senescence marker protein SMP-30 - rat
N:Alternate names: regucalcin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C:Accession: S34588; S34589; S27203
R:Shimokawa, N.; Yamaguchi, M.
FEBS Lett. 327, 251-255, 1993
A:Title: Molecular cloning and sequencing of the cDNA coding for a calcium-binding protein SMP-30 from rat
A:Reference number: S34588; MUID:93351639; PMID:8348951
A:Accession: S34588
A:Molecule type: mRNA
A:Residues: 1-299 <SHJ>
A:Cross-references: GB:D38467; GB:D14327; GB:D16386; NID:g600378; PIDN:BAAG07490.1; PI
A:Accession: S34589
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 84-299 <SH2>
R:Fujita, T.; Shirasawa, T.; Uchida, K.; Maruyama, N.
Biochim. Biophys. Acta 1132, 297-305, 1992
A:Title: Isolation of cDNA clone encoding rat senescence marker protein-30
A:Reference number: S27203; MUID:93041931; PMID:1420310
A:Accession: S27203
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-147, 'D', 149-299 <FUJ>
A:Cross-references: EMBL:X69021; NID:g57254; PIDN:CAA48786.1; PID:g57255
C:Superfamily: senescence marker protein-30
C:Keywords: calcium binding

Query Match 28.1%; Score 454.5; DB 2; Length 299;
Best Local Similarity 37.5%; Pred. No. 3.4e-29;
Matches 111; Conservative 41; Mismatches 121; Indels 23; Gaps 7;

QY 14 YTVGEGPHMDHETQTLFYFVDVTEKTFHKYVPSQKKTFFCKV--DKLVSF 73
DB 14 YRCGESPVWEEASQCLLFVDIPSKTVCRWDSISNRVQRVGDVAPVSSV--ALRQSGGYVA 71
QY 74 SLEREIALTWDGVSAAPTSTEAIVNVPEPHIKNNRNLNDGKADPLGNLWGTGTMAIDAGLPV 133
DB 72 TIGTKFCALNWE-----QSVFLAMVDEDDKNNRFDNKGKVDPAGRYFAGTMAEE---T 122
QY 134 GPVT-----GSLYHLGADKKVKHESNIAIANGLAWSNDLKKMYIDSGRRRVDEYDYDA 188

Db 123 APAVLERHQGLSLFDPHDSVKKYFNQVDISNGLDMSLDHDKFYIYIDSLSTVTVDAFDYDL 182
Qy 189 STLISNORPLTFTEKHEVPGYDQQTIDEENGLWVAVFQGORIHKISTQOPEV--LLDT 246
Db 183 PTGQISNRRYVYKNEKDE--QIPDGMCIDVSGKLWVACVNGRVRIL---DPETGKRLQT 237
Qy 247 VKIPDPQVTSVAFGPNDELHVTSGAGLQDSSLSKLVNGHYVVTGLGVKQGA 302
Db 238 VKLPVDPTTSCFCGKDYSEMYVTVCARDGMSAEGLLRQPDAGNIFKITGLGVKQGA 293
RESULT 5
S60035
senescence marker protein 30 - human
N:Alternate names: regucalcin
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: S60035; 152491
R:Fujita, T.; Mandel, J.L.; Shirasawa, T.; Hino, O.; Shirai, T.; Maruyama, N.
Biochim. Biophys. Acta 1263, 249-252, 1995
A:Title: Isolation of cDNA clone encoding human homologue of senescence marker protein-3
A:Reference number: 152491; MUID:96004897; PMID:7548213
A:Accession: S60035
A:Molecule type: mRNA
A:Residues: 1-299 <FU0>
A:Cross-references: EMBL:D31815; NID:g1072311; PIDN:BAA06602.1; PID:g1072312
C:Genetics:
A:Gene: GDB:RGN; RC: SMP30
A:Cross-references: GDB:9955055
A:Map position: Xp11.3:Xp11.2
C:Superfamily: senescence marker protein-30
C:Keywords: calcium binding

Query Match 27.8%; Score 448.5; DB 1; Length 299;
Best Local Similarity 36.9%; Pred. No. 1e-28;
Matches 110; Conservative 42; Mismatches 113; Indels 33; Gaps 8;
Qy 17 GEGPHWDHETOTLYFVDVTEKTFHKYVPSOKKY---TFCKVDKLVSTFIPLAG-----SPG 69
Db 17 GESPWVEVSNLLFVD-----IPAKVKCRWDSFTKQVQRVTMDAPVSSVALRQSG 67
Qy 70 RFVVSLEIEAILTWGDSVSAPTSIEAIVNVEPHKNNRLNDGKADPLGNLWTGTMADA 129
Db 68 GYVATIGTKFCCALMKWESAV-----VLATVDNKKNNRNDGKYDPAGRYFAGTMAEE- 121
Qy 130 GLPVGPPV-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSQGRKRVDEY 184
Db 122 ---TAPAVLERHQGLYSLFDPHVKYFDQVDVDSGLDWSLDHDKIFYIYIDSLSYSDAF 178
Qy 185 DYDASTLSISNORPLTFTEKHEVPGYDQQTIDEENGLWVAVFQGORIHKISTQOPEVLL 244
Db 179 DYDLQTGQISNRRSYVYKLEKEE--QIPDGMCIDAEGLWVACVNGRVRILDPVTGK-RL 235
Qy 245 DTVKIPDPQVTSVAFGPNDELHVTSGAGLQDSSLSKLVNGHYVVTGLGVKQGA 302
Db 236 QTVKLPVDKTTSCFCGKDYSEMYVTVCARDGMSAEGLLRQPDAGNIFKITGLGVKQGA 293

RESULT 6
E95309
probable calcium binding transcription regulatory protein [imported] - Sinorhizobium meli
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E95309
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: E95309
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-569 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65039.1; PID:g14523470; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSYMA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0717
A:Genome: plasmid

Query Match 26.2%; Score 423.5; DB 2; Length 569;
Best Local Similarity 34.6%; Pred. No. 2.7e-26;
Matches 100; Conservative 46; Mismatches 102; Indels 41; Gaps 7;
Qy 16 VEGEPHWDHETOTLYFVDVTEKTFHKYVPSOKKYTFCKVDKLVSTFIPLAGSPGRFVVS 75
Db 293 LGEAPVWVEREKRLWVDLHPAVHRFDPVTCKNESCNAVLSAVLP-----TR 342
Qy 76 EREIALLTWGDSV-----AAPTSTIEAIVNVEPHKNNRLNDGKADPLGNLWTG 123
Db 343 NEGLIVASQDGVVEHFDFDRGDFNPPAEP-----EPGLPENRLNDKAVDPSGRLWVG 393
Qy 124 TMAIDAGLPVGPVTSGLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSQGRRVDE 183
Db 394 SMRLDVSRLP-----TGSLYRLTSAGEVTRAGSGFTVANGLAWSNDSSTFYFVDVTPGIIYA 449
Qy 184 YDYDASTLSISNORPLTFTEKHEVPGYDQQTIDEENGLWVAVFQGORIHKISTQOPEVL 243
Db 450 YDFDAREGSIANRRVFTVP--EAGSRPDGLAVDADGGVCAIWDGWR---VNRYPDGR 504
Qy 244 LD-TVKIPDPQVTSVAFGPNDELHVTSGAGLQDSSLSKLVNGHYV 291
Db 505 LDRVAVLPVPRPTSVAFGGDELATLFTSARTLRPLASTLTPAELSGGIF 553

RESULT 7
A13216
hypothetical protein Atu5464 [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: A13216
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: A13216
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-560 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL46151.1; PID:g17743920; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5464
A:Genome: plasmid

Query Match 25.4%; Score 409.5; DB 2; Length 560;
Best Local Similarity 33.1%; Pred. No. 3.6e-25;
Matches 99; Conservative 57; Mismatches 110; Indels 33; Gaps 9;
Qy 16 VEGEPHWDHETOTLYFVDVTEKTFHKYVPSOKKYTFCKVDKLVSTFIPLAGSPGRFVVS 75
Db 284 LGESPWVEGEGDALYWDVLLHPAVHRFDPATGRNETCEGTGKLVSAVPTG--GRLLVAS 341
Qy 76 EREIALLTWGDSV---SAAPTSTIEAIVNVEPHKNNRLNDGKADPLGNLWTGTMADAGLP 132

```
Db 342 Q-----DGVEMDLFASGLRTPFVSPAGIANRUNDACKGPDGAIWVGSMRIDASKP 393
QY 133 VGPVTSGLYHLGADKKVKMHESNTAIAAGLAWSLNDLKKMYIDSGKRVDYDASTLS 192
Db 394 ---TGALYRINANGASERKEGGIIVNSGLWSPDGRFTYFVDPVPLGIIHAYDCDPATGA 449
QY 193 ISNQRLPFTFEKHEVP---GYPDGQTIIDEGNLWAVFQGORIITKISTQOPEVLID-TVK 248
Db 450 LSQRR-----EFARIPVADGRPDGLAVDAEGVWCWVRR--YLPNGKLDQVID 501
QY 249 IPDPQVTSVAFGPNLDLDELHVTAGLQDSSLSLKVHNGHYR-----VTGLGKVGFRAG 303
Db 502 MPVPRPSIATGPPDLSTLFTISARTLPLASTLADAPLSGLGSCRPQISGARIAMFEG 560

RESULT 8
A70047
RNA polymerase homolog yvE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A70047
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chedid, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tanakoshi, A.; Tanaka, H.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A70047
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-292 <KUN>
A:Cross-references: GB:299120; GB:299121; GB:AL009126; NID:92635827; PID:92635833; NID:92635834
A:Experimental source: strain 168
C:Genetics:
A:Gene: yvE
C:Superfamily: senescence marker protein-30
```

```
Query Match 23.3%; Score 376; DB 1; Length 292;
Best Local Similarity 31.5%; Pred. No. 7.2e-23;
Matches 85; Conservative 52; Mismatches 105; Indels 28; Gaps 7;

QY 16 VGGSPHMDHETOTLYFVDVVEKTFHKYVPSQKKTFCVKDLVSVFTIPLAGSPGRFVWSL 75
Db 13 IGGPPLWDEENGLYWDILGSELHIFDPEE-----KINRSIKF-----KSFVTL 58

QY 76 ER----EIALTWG---VSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAID 128
Db 59 AKYSKDELIMTKDGFYHLRDSLEKIKQPKDMHESLRFNDAKDPYGLRWAGTTSME 118

QY 129 AGLPVGTVGSLYHLGADKKVKMHESNTAIAAGLAWSLNDLKKMYIDSGKRVDYDVA 188
Db 119 G----EQKQASLYRLNDGSLVKIKDQVSTNSGLDMDRNLMLYIDTPTQELVIRYSYDP 174

QY 189 STLSISNORLFFFEKHEVPGPDQTIIDEGNLWAVFQGORIITKISTQOPEVLIDTVK 248
Db 175 QSGDVNSPEVYFDQSD--GLPDGWTIDQNGMLWALFGSRVVIDFQKKE-INSIS 231

QY 249 IPDPQVTSVAFGPNLDLDELHVTAGLQDSSLSLKVHNGHYR-----VTGLGKVGFRAG 303
Db 232 VPAKYVTCGAFGGRDLKTLTYITTAQMT 261
```

```
RESULT 9
B90445
regucalcin homolog [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: B90445
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: B90445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-285 <KUR>
A:Cross-references: GB:AE006641; NID:913816023; PIDN:AAK42817.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO2705
C:Superfamily: senescence marker protein-30
```

```
Query Match 22.6%; Score 365.5; DB 2; Length 285;
Best Local Similarity 31.6%; Pred. No. 4.9e-22;
Matches 98; Conservative 55; Mismatches 106; Indels 51; Gaps 12;

QY 13 KYTVGEGPHMDHETQTLFYVD-----TVEKTFHKYVPSQKKTFCVKVD--KLV 58
Db 12 KATLGEPPVYDKELNKLFWVDIECKRIIVNDLNTGTETVEVE-MPDLIS-SLCVIDDKRVI 69

QY 59 SFTIPLAGSPGRFVSVLERETAITWDGVSAPTSIAIVNVEPHIKNNRLNDGKADPLG 118
Db 70 ATI-----RHGFYVNLDR-----GEQKIAEVETDMESNRFNDGKCDKLG 110

QY 119 NLWTGTWAIADAGLPVPGTGSGLYHLGADKKVKMHESNTAIAAGLAWSLNDLKKMYIDSGK 178
Db 111 RYWAGTMNNRKP-----TGSFYKLDTKITKILE-GLTVSNGLWDPDNTQMFLLIDSPV 165

QY 179 RRVDEYDYDASTLSISNORLFFTEKHEVPGPDQTIIDEGNLWAVFQGORIITKISTQ 238
Db 166 RKVFVDFDLNKGELYNRRVAVDFGNE--PCNPDMGMAVDEEGYIWAHMGCGKYSRWNPK 223

QY 239 QPEVLIDTVKTPDQVTSVAFGPNLDLDELHVTAGLQDSSLSLKVHNGHYRVTGLGV 298
Db 224 TGKKLFE-IKVPATYVTSVTGTQELDLQFLITTAGKSQDP-----LAGKTF-TTKANV 274

QY 299 KGFAGVKVKL 308
Db 275 RGLQNFREKI 284
```

```
RESULT 10
A87351
Smp-30/Cyrl family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: A87351
R:Nierman, M.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87351
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <STO>
A:Cross-references: GB:AE005673; NID:913422069; PIDN:AAK22805.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0820
C:Superfamily: senescence marker protein-30
```

```
Query Match 20.5%; Score 331; DB 2; Length 289;
Best Local Similarity 33.3%; Pred. No. 3.1e-19;
```

Matches	94; Conservative	47; Mismatches	121; Indels	20; Gaps	10;
QY	13 KYTVGEGPHDHTQTLYFYDVTVEKTHPKVPSQ--KKYTECKVDKLVSFIIPLAGSPGRF	71			
Db	12 KATLGEGPIWHGD--TLWFDVIQKQTHNYHPATGERSEFDAPDQ-VTFLLAPIVGATG-F	67			
QY	72 VVSLEREIALITWDGVSAAPSTSEAIYVNEPHIKNNRLNDGKADPLGNLWLTGTWDAIDL	131			
Db	68 VVGLKGTGIHRF-----HPATGESLLLEVEDAALNNRPNDATVDAQGRLWFGTWHGDE--	119			
QY	132 PVGPFVTSGLYHLGADKKVKMHESIAIANGLAWSNDLKMYIIDSGRKRRVDEYDYDASTL	191			
Db	120 --ENNSSGLYRMDLTGVARM--DRDICITNGPCVSPDGKTFVHTDTLEKTIIVAFDLAEDGL	176			
QY	192 SISNQRLPWFKEHVEVPGYDQGTIDEGNLVAVFQGORIIKISTQOPEVLLDVTVKIPD	251			
Db	177 -LSNKRVYQFALGD-DVYPDGVSVDSEGYLWTALMGWFGAVRSEPOGDV--TRIELPA	232			
QY	252 PQVTSVAFGPNLDLHVTSAGLQDSSLDKSLVNGHVYRV	293			
Db	233 PNVTKPCGCGPDLKTLFTTARKGLSDETLAOYVPLAGSVFV	274			

RESULT 11
A87473
Smp-30/Cgrl family protein [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C/Accession: A87473
B.; Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heid
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fra
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:211173698; PMID:11259647
A/Accession: A87473
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-293 <SPQ>
A/Cross-references: GB:AE005673; NID:gl3423239; PIDN:AAK23781.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC1805
C/Superfamily: senescence marker protein-30

		Query Match	19.5%; Score 315; DB 2; Length 293;
		Best Local Similarity	30.9%; Pred. No. 6.2e-18;
		Matches	93; Conservative 51; Mismatches 115; Indels 42; Gaps 13;
QY	16	VGEPSHWDHTQTLYFYDVTVEKTHKVPSOKKYTCFKVDVKLVFSFILPAGSPGRFFVSL	75
Db	15	LGSPLWADAGVLVIYWDSPAATWRYPDFTSEQRSTAPAKPGISVV--LGRPGLIAGL	72
QY	76	EREAIILTGDGVSAAPTSI-EAIVNVEPHIKNNRLNDGKADPLGNLMWTGTMAI--DAGLP	132
Db	73	ADGYRVOLDTGAFTPIALPDTLAPIE-----RENDGKAIDRGFEVYTGTMAMHNETGR-	125
QY	133	VGPVTSGLYHLGADKKVMYESN-IATANGLAWSNDLKKNYYIDSRRVRDEYDYASTL	191
Db	126	----IGKLRYFSAGAWAEVLPTPEIEIANSTCFSPSGDTLYFADSLRHVMWFASYPDKTG	181
QY	192	SISNRPLTFTEKHKEVPGY---PDGQITDERGNLWAVFGORIIKITSTQOPEVLDD-TV	247
Db	182	AVEKRDFP-----DTTGFNSAPDGATVDAGHIWLALVQAOKLRIS---PDGRLDRW	233
QY	248	KIPDPQVTSVAFGPNLDLHVTSAGLQLDPS-----SLDSKLVNGHHVYRTGLGVKGF	301
Db	234	ESPAPFCSCPAFGEDLLIYVTS- ---ISDSGGRLKTDVDAS---GRLFMAFHGLGVRSI	286
QY	302	A 302	
Db	287	A 287	

RESULT 12
G90485 conserved
C:Species
C:Date: 2
C:Accessi
R:She, Q.
Jong, I.
arrett, R
submitted
A:Descrip
A:Referen
A:Accessi
A:Status:
A:Molecu
A:Residue
A:Cross-r
C:Genetic
A:Gene: S
C:Superfa

Query Match Best Local Matches
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db

RESULT 13
D95265
probable regucalcin [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSyC
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95265
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: D95265
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <KUR>
A:Cross-references: GB:AE006469; PID:AAK64686.1; PID:g14523085; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSyMA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:

A:Gene: Sma0060
A:Genome: plasmid
C:Superfamily: senescence marker protein-30

Query Match 17.9%; Score 288.5; DB 2; Length 311;
Best Local Similarity 28.18; Pred. No. 9.4e-16;
Matches 85; Conservative 47; Mismatches 128; Indels 43; Gaps 11;
Db 12 KDIVGESILWGDDEKALVWVDIVCKRIHRLPEGRHDTWPTDFVTSI----GMRKDG 67
Qy 13 KYTVGEGPHMDHETQTLFYVDVTEKFKYVPSQKTYFCKVKLVSFIIPLAG--SPGR 70
Db 71 FVSLERERIALTWGVSAAPTSTEAIVNVEPHKNNRLNDGKADPLGNLWTGTMA--ID 128
Db 68 FIVGLSRVNCVLTWDG-----PFEFAMPEPDLLENRNEGRVAPDGSFWATMQSNLD 121
Qy 129 AG--LPVGPVTGSLYHLGADKKV-KMHESNIAIANGLAWSNLDKMKYIYDSGKRRVDEY 184
Db 122 AGGSPKMDROSGAVRIDPTGHVSQLTPEYGITNTMGWTRD-NREFFADTLANEIYMF 180
Qy 185 DYDASTLSISNORPLFT-FEKHEVPGYPDGOTIDEENLW-----VAVFGQRIIK 234
Db 181 DCDLAARRIDNRRTIVAGFAR---GLPDGSCLDADDRLMNCRVAGGAAGVAFDG----- 231
Qy 235 ISTQOPEVLLDTVKTPDQVTSVAFGGPNLDELHVTSAAGLQDSSLDKSLVNGHVYRVT 294
Db 232 -----AGRLMHLIELPASWPTSCFTGGPVLTSLVTSARFTWTDHMDHMPLEGGLFAVE 286
Qy 295 GLG 297
Db 287 GVG 289

RESULT 14

AB3071
calcium-binding protein, regucalcin [imported] - Agrobacterium tumefaciens (strain C58,
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AB3071
R:Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB3071
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL44984.1; PID:g17742642; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4190
A:Map position: linear chromosome
C:Superfamily: senescence marker protein-30

Query Match 16.9%; Score 273; DB 2; Length 293;
Best Local Similarity 31.1%; Pred. No. 1.6e-14;
Matches 82; Conservative 39; Mismatches 113; Indels 30; Gaps 10;
Qy 16 VGEQPHDHTOTLYFVDVTEKFKYVPSQKTYFCKVKLVSFIIPL-AGSPG----- 69
Db 21 VGESPTWERTGDLWFVDILAPAFCLSPS-----GKLOREDMPAQIGCLGLCQSN 71
Qy 70 RFVVSLEIRETAILTWGVSAAPTSTEAIVNVEPHKNNRLNDGKADPLGNLWTGTMAIDA 129
Db 72 LIVAAKLTGVHLL-----NPQNGELELLCDPDDEGRADSLNDGKVGPDGHEFWVGTDR- 125
Qy 130 GLPVGVPVTSYHLGADKKV-KMHESNIAIANGLAWSNLDKMKYIYDSGKRRVDEYDYDA 188
Db 126 DVPTG--NARLYRVSSSGHVERFDGDMLTNSGLAWSPDETMYHSDSGLLLQVDFDV 183

Qy 189 STLSISNORPLFTFEKHEVPGYPDGOTIDEENLWVAVFQGORIKISTQOPE-VLLDVT 247
Db 184 QTRGLGPAKRLHDFQDE--GRPDGAATDSEGCYSAGVQAGRLNRFT---PDGELFEIY 238
Qy 248 KIPDPQVTSVAFGGPNLDELHVT 271
Db 239 KMPFGKPTMPCFGGPELKTLYVTS 262

RESULT 15

G98215
senescence marker protein-30 (AB033368) [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 15-Feb-2002
C:Accession: G98215
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G98215
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89249.1; PID:g15159075; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_1336
A:Map position: linear chromosome
C:Superfamily: senescence marker protein-30

Query Match 16.9%; Score 273; DB 2; Length 322;
Best Local Similarity 31.1%; Pred. No. 1.8e-14;
Matches 82; Conservative 39; Mismatches 113; Indels 30; Gaps 10;
Qy 16 VGEQPHDHTOTLYFVDVTEKFKYVPSQKTYFCKVKLVSFIIPL-AGSPG----- 69
Db 50 VGESPTWERTGDLWFVDILAPAFCLSPS-----GKLOREDMPAQIGCLGLCQSN 100
Qy 70 RFVVSLEIRETAILTWGVSAAPTSTEAIVNVEPHKNNRLNDGKADPLGNLWTGTMAIDA 129
Db 101 LIVAAKLTGVHLL-----NPQNGELELLCDPDDEGRADSLNDGKVGPDGHEFWVGTDR- 154
Qy 130 GLPVGVPVTSYHLGADKKV-KMHESNIAIANGLAWSNLDKMKYIYDSGKRRVDEYDYDA 188
Db 155 DVPTG--NARLYRVSSSGHVERFDGDMLTNSGLAWSPDETMYHSDSGLLLQVDFDV 212
Qy 189 STLSISNORPLFTFEKHEVPGYPDGOTIDEENLWVAVFQGORIKISTQOPE-VLLDVT 247
Db 213 QTRGLGPAKRLHDFQDE--GRPDGAATDSEGCYSAGVQAGRLNRFT---PDGELFEIY 267
Qy 248 KIPDPQVTSVAFGGPNLDELHVT 271
Db 268 KMPFGKPTMPCFGGPELKTLYVTS 291

Search completed: August 1, 2003, 14:35:00
Job time : 41 secs

192 SISNORPLTFTEKHEVPGYPDGOTIDEEGNLWVAVFOGORIIKISTOOPEV--LLDPTVKI 249


```

Db 186 QISNRRIYKLEKEE--QIPDGMCIDTEGKLVWACYNGGRVIRL---DPETGKRLOTVKL 240
QY 250 PDQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLGVKRG-----FAG 303
Db 241 PVDKTTSCCGFGKDYSEMYVTCARDGLDPSLSRQPEAGGIFKITGLGVKGIPIPPYSVAG 299

```

RESULT 2

```

SM30_MOUSE
ID SM30_MOUSE STANDARD; PRT; 299 AA.
AC G64374; Q60944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN RGN OR SMP30.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=96328264; PubMed=8765750;
RA Fujita T., Shirasawa T., Maruyama N.;
RT "Isolation and characterization of genomic and cDNA clones encoding
RT mouse senescence marker protein-30 (SMP30).";
RL Biochim. Biophys. Acta 1308:49-57(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97422495; PubMed=9278463;
RA Murata T., Yamaguchi M.;
RT "Molecular cloning of the cDNA coding for regucalcin and its mRNA
RT expression in mouse liver: the expression is stimulated by calcium
RT administration.";
RL Mol. Cell. Biochem. 173:127-133(1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMMATIC ACTIVITY
CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC CALCIUM SIGNALING IN THE AGED LIVER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: MAINLY PRESENT IN THE LIVER. WEAK EXPRESSION
CC WAS FOUND IN THE BRAIN AND LUNG, NOT FOUND IN THE KIDNEY.
CC -!- DEVELOPMENTAL STAGE: PROTEIN AMOUNTS IN LIVER DECREASE
CC SIGNIFICANTLY WITH AGE.
CC -!- SIMILARITY: BELONGS TO THE SMP-30 / CGRI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28937; AAC52721.1; -
DR EMBL: U32170; NAD03478.1; -
DR EMBL: D86217; BAA13046.1; -
DR SWISS-2DPAGE: Q64374; MOUSE.
DR MGD; MGI:108024; Rgn.
KW Calcium-binding.
SQ SEQUENCE 299 AA; 33407 MW; DAD55EF618311977 CRC64;

```

```

Query Match 28.3%; Score 457.5; DB 1; Length 299;
Best Local Similarity 37.8%; Pred. No. 3.8e-29;
Matches 112; Conservative 40; Mismatches 121; Indels 23; Gaps 7;

QY 14 YTVGEPHMDHETOTLYFYDVTVEKTFHKYVPSQKYYTCKVDKLVSFIIPLAGSPGRFVY 73
Db 14 YRCGSPVWEASQSLFLFDIPSKICRWDTVSNQVRVADPVSIV--ALRQLGGYVA 71
QY 74 SLEREITAILTDGVSAAPTSIETAIENVNVEPHIKNNRLNDGKADPLNLTGTMAIDAGLPV 133

```

```

Db 72 TGTGTCFALNWNEN-----QSVFVLAWVDEDKNNRNDGKVDPAGRYPAGTMAEE---T 122
QY 134 GPVT-----GSYLHGLADKKVKMHESNIAIANGLAWSNDLKKMYIYIDSGKRRVDEYDIDA 188
Db 123 APAVLERHQSGLSYLPDPHSVAKKYFDQVDISNGLDWSLDHKIFYIYIDSLSYTVDAPFDYL 182
QY 189 STLSISNORPLTFEKKHEVPYGDGTIDEENLWAVFQGORIIKISTOOPEV--LLDT 246
Db 183 QTGQISNRRIVYKMEKDE--QIPDGMCIDAEGLKLVWACYNGGRVIRL---DETGKRLOT 237
QY 247 VKIPDPQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLGVKGA 302
Db 238 VKLPVDKTTSCCGFGKDYSEMYVTCARDGLNAGELLRQDPDAGNIPKITGLGVKGA 293

```

RESULT 3

```

SM30_BOVIN
ID SM30_BOVIN STANDARD; PRT; 299 AA.
AC Q9TJ75;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN RGN OR SMP30.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20351777; PubMed=10891565;
RA Misawa H., Yamaguchi M.;
RT "The gene of Ca2+-binding protein regucalcin is highly conserved in
RT vertebrate species.";
RL Int. J. Mol. Med. 6:191-196(2000).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMMATIC ACTIVITY
CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC CALCIUM SIGNALING IN THE AGED LIVER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMP-30 / CGRI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB035446; BAA88080.1; -
KW Calcium-binding.
SQ SEQUENCE 299 AA; 33308 MW; 4258F7CDB3CCB575 CRC64;

```

```

Query Match 28.1%; Score 454.5; DB 1; Length 299;
Best Local Similarity 38.1%; Pred. No. 6.6e-29;
Matches 111; Conservative 38; Mismatches 119; Indels 23; Gaps 7;

QY 17 GEGPHWDHETOTLYFYDVTVEKTFHKYVPSQKYYTCKVDKLVSFIIPLAGSPGRFVVSLE 76
Db 17 GESPVWEASQSLFLFDIPAKKVKCRWDSLSKQVRVTDAPVSV--ALRQSGGYVATVG 74
QY 77 REIAILTDGVSAAPTSIETAIENVNVEPHIKNNRLNDGKADPLNLTGTMAIDAGLPVGPV 136
Db 75 TKFCALNWDGSAV-----VLATVDKKEKKNNRNDGKVDPAGRYFAGTMAEE---TAPA 125
QY 137 T-----GSYLHGLADKKVKMHESNIAIANGLAWSNDLKKMYIYIDSGKRRVDEYDASTL 191
Db 126 VLERRQSGLSYLPDPHSHVKEYFDQVDISNGLDWSMDHKIFYIYIDSLSYSDVAFDYDLOTG 185
QY 192 SISNORPLTFEKKHEVPYGDGTIDEENLWAVFQGORIIKISTOOPEV--LLDTVKI 249

```



```

Db 186 KISNRSTVAFKEE--QIPDGMCDIVVEGKLWVACYNNGRVIRL---DPETGRRLQTVKL 240
QY 250 PDQVTSVAFGGPNLDLHVTSAGLQDSSLDKSLVNGHVYRVTVGLGVKG 300
Db 241 PVDKTTSCCEGGKDYSEMYVTCARDGLDCKGLLQOPEAGGIFKTVGLGVKG 291

RESULT 4
SM30_RAT
ID SM30_RAT STANDARD; PRT; 299 AA.
AC Q03336; Q63496;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN RGN OR SMP30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=93041931; PubMed=1420310;
RA Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT "Isolation of cDNA clone encoding rat senescence marker protein-30
RT (SMP30) and its tissue distribution.";
RL Biochim. Biophys. Acta 1132:297-305(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=93351639; PubMed=8348951;
RA Shimokawa N., Yamaguchi M.;
RT "Molecular cloning and sequencing of the cDNA coding for a calcium-
RT binding protein regucalcin from rat liver.";
RL FEBS Lett. 327:251-255(1993).
RN [3]
RP SEQUENCE OF 117-299 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=96150990; PubMed=8569761;
RA Shimokawa N., Matsuda Y., Yamaguchi M.;
RT "Genomic cloning and chromosomal assignment of rat regucalcin gene.";
RL Mol. Cell. Biochem. 151:157-163(1995).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=96386712; PubMed=8794449;
RA Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT "Gene regulation of senescence marker protein-30 (SMP30): coordinated
RT up-regulation with tissue maturation and gradual down-regulation
RT with aging.";
RL Mech. Ageing Dev. 87:219-229(1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMIC ACTIVITY
CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC CALCIUM SIGNALING IN THE AGED LIVER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HEPATOCYTES AND RENAL PROXIMAL TUBULAR
CC EPITHELIUM.
CC -1- DEVELOPMENTAL STAGE: IN LIVER, THE FIRST PEAK OF EXPRESSION WAS
CC FOUND IN 5-DAY-OLD NEONATES. EXPRESSION INCREASES FROM DAY 7 AND
CC REACHES A PLATEAU AT DAY 10. 3-6.5 MONTH-OLD ADULTS EXPRESS ABOUT A
CC THIRD THE AMOUNT OF NEONATES LEVEL. IN KIDNEY, EXPRESSION
CC INCREASES FROM DAY 21 AND REACHES A MAXIMAL LEVEL AT DAY 35,
CC REMAINS HIGH UNTIL 3 MONTHS OF AGE.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC
CC EMBL; X69021; CAA48786.1; -
CC DR EMBL; D38467; BAA07490.1; -
CC DR EMBL; D31662; BAA06507.1; -
CC DR PIR; S27203; S27203.
CC KW Calcium-Binding.
CC FT CONFLICT 148 148 N -> D (IN REF. 1).
CC SQ SEQUENCE 299 AA; 33389 MW; E3CF6D3FFCAE4E98 CRC64;

Query Match 28.1%; Score 454.5; DB 1; Length 299;
Best Local Similarity 37.5%; Pred. No. 6.6e-29;
Matches 111; Conservative 41; Mismatches 121; Indels 23; Gaps 7;

QY 14 YTVGEGPHWDHETQTLTYFVDTVEKTFKHYVPKQKYYTFCVKDKLVSLFIIPLAGSPGRFVV 73
Db 14 YRCGESPVWEEASKLLFVDIPSKTVCRWDSISNRVQGVGDAPVSSV--ALRQSGGYVA 71
QY 74 SLERETAILTWGVSAAPTSEIAIVNVEPHIKNNRLNDGKADPLGNLWGTMTAIDAGLPV 133
Db 72 TIGTKFCALNWE-----DQSVFILAMVDEDKNNRDNKRVDPAGRYFAGTMAEE----T 122
QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKMYYIDSGRRRVDEYDYDA 188
Db 123 APAVLERHQGLSLYLPDPHSHVKYTFNQVDISNGLDWLSLHKIFYYIDLSYTVDAFDYDL 182
QY 189 STLSISNRPLTFEKEHEVPGYDQGTIDEEGNLWAVFQGRIRIKISTQQPEV--LLDT 246
Db 183 PTGQISNRRTVYKMEKDE--QIPDGMCDIVVEGKLWVACYNNGRVIRL---DPETGRRLQTV 237
QY 247 VKIPDPQVTSVAFGGPNLDLHVTSAGLQDSSLDKSLVNGHVYRVTVGLGVKGFA 302
Db 238 VKLPVDKTTSCCEGGKDYSEMYVTCARDGMSAEGLLRQPDAGNIFKTVGLGVKGIA 293

RESULT 5
SM30_HUMAN
ID SM30_HUMAN STANDARD; PRT; 299 AA.
AC Q15493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN RGN OR SMP30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96004897; PubMed=7548213;
RA Fujita T., Mandel J.L., Shirasawa T., Hino O., Shirai T., Maruyama N.;
RT "Isolation of cDNA clone encoding human homologue of senescence
RT marker protein-30 (SMP30) and its location on the X chromosome.";
RL Biochim. Biophys. Acta 1263:249-252(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=10677570;
RA Misawa H., Yamaguchi M.;
RT "Transcript heterogeneity of the human gene for Ca2+-binding protein
RT regucalcin.";
RL Int. J. Mol. Med. 5:283-287(2000).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMIC ACTIVITY
CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC CALCIUM SIGNALING IN THE AGED LIVER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; D31815; BAA06602.1; --
DR EMBL; AB028125; BAA78693.1; --
DR EMBL; AB032064; BAA84082.1; --
DR Genbank; HGNC:9989; RGN.
DR MIM; 300212; --
KW Calcium-binding.
SQ SEQUENCE 299 AA; 33253 MW; 95BALC73B7B7635 CRC64;

Query Match 27.8%; Score 448.5; DB 1; Length 299;
Best Local Similarity 36.9%; Pred. No. 2e-28;
Matches 110; Conservative 42; Mismatches 113; Indels 33; Gaps 8;

QY 17 GEGPHDHEOTLYFVDVTEKTHKVPQSKY---TFCKVDKLVFSFIPLAG---SPG 69
DB 17 GESPVEEVSLLFVD-----IPAKKVCNDSFTKQVQVVTMDAPVSSVALRQSG 67
QY 70 FVVSLEIRETAILTDCVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDA 129
DB 68 GYVATICFKCALNWKQSAV-----VLATVDNDKNNRENDGKVDPAGRYFAGTMAEE- 121
QY 130 GLPVPVY-----GSLYHLGADKKVKWHESNIAIANGLAWSNDLKKWYIDSKRRVDEY 184
DB 122 ---TAPAVLERHQGLASLFPDHHKKYFQVQVDSNGLWSLHKIFYYIDSLSYSDAF 178
QY 185 DYDASTLSISNORPLETFEKEHVPYDPGOTIDEGNLVAVFOGQRIIKISTQOQPEVLL 244
DB 179 DYDQGISNRRSVYKLEKEE--QIDGMCIDAEGLWVACVNGRVRDLDPVTGK-RL 235
QY 245 DTVKIPQVTSVAFGGPNLDLHVTSAGLQLDSSLDKSLVNGHYVTVGLGVKGF 302
DB 236 QTVKLPVDKTSCEGCKNYSMTVCARDGMDEGLLRQPEAGGIFKTLGLGVKGIA 293

RESULT 6
Y3IK_SULAC
ID Y3IK_SULAC STANDARD; PRT; 275 AA.
AC P46218;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 31.5 kDa protein.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OX NCBI_TaxID=2285;
RN
RP SEQUENCE FROM N.A.
RA Durovic P.V., Potter S., Dennis P.P.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: BELONGS TO THE SMP-30 / CGRI FAMILY.
CC -! CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A RNA POLYMERASE
CC SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U05664; AAA73413.1; --
SQ SEQUENCE 275 AA; 31544 MW; D949D5BF83EBE582 CRC64;

Query Match 24.2%; Score 391.5; DB 1; Length 275;
Best Local Similarity 33.0%; Pred. No. 5.7e-24;
Matches 98; Conservative 51; Mismatches 97; Indels 51; Gaps 10;

QY 18 GEPHWDHETOTLYFVDVTEKTHKVPQSKYKTFCKVDKLVFSFIIP-----LAGSPGR 70
DB 13 EGPIWAY--NSLYFVDIPKGLHNL-LKEDGTHVWVKFPTTYVSSLOPTVRGGIIVTAGNGF 69
QY 71 FVVSLEIRETAIL-----TWDGVSAAPTSTIEAIVNVEPHIKNNRLNDGKADPLGNLTGTMA 126
DB 70 YLVKDKDQISLLEYKDW-----SRNRFNDGKDKQMGRYWIGTMN 110
QY 127 IDAGLPVGVGTSLYHLGADKKVKWHESNIAIANGLAWSNDLKKWYIDSKRRVDEYDY 186
DB 111 LEEKYP---TGGLFVLDLDMKFRRLVTDVITNSLAWSLDNKLYYIDSPTRKTFKPKF 166
QY 187 DASTLSISNORPLETFEKEHVPYDPGOTIDEGNLVAVFOGQRIIKISTQOQPEVLLDT 246
DB 167 DIERGDISOREVLIDLKEYE--GVDPGCTIDSEGNLWALYGGAVLRIDVEKRVIOE- 223
QY 247 VKIPQVTSVAFGGPNLDLHVTSAGLQLDSSLDKSLVNGHYV--RVTVGLGVKGF 301
DB 224 LRLPAPRVTSVIFGGNSMDTLFTITANDHPD-----GGFVYSERVDVKGVEY 271

RESULT 7
YVRE_BACSU
ID YVRE_BACSU STANDARD; PRT; 292 AA.
AC Q34940;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yvre.
OS YVRE.
GN Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98304083; PubMed=9639930;
RA Wipat A., Brignell C.S., Guy J.B., Rose M., Emmerson P.T.,
RA Harwood C.R.;
RA "The yvsa-yvqa (293 degrees - 289 degrees) region of the Bacillus
RT subtilis chromosome containing genes involved in metal ion uptake and
RL a putative sigma factor."
RN Microbiology 144:1593-1600(1998).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Deniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrarini E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Madigic C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

CC	EMBL; 235922; CAA84996.1; -
DR	EMBL; 246260; CAA86396.1; -

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AB001684; BAA57886.1; -
CC HSSP; P07157; 1A1P.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU.Cterm.
DR InterPro: IPR004161; EFTU.D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR

```
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNCT.
DR TIGRFAMS; TIGR00231; smallGTP; 1.
DR TIGRFAMS; TIGR00485; EF-TU; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; Chloroplast;
KW GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 409 AA; 44911 MW; 0183143E5E117999 CRC64;

Query Match 5.9%; Score 94.5; DB 1; Length 409;
Best Local Similarity 19.6%; Pred. No. 3.2;
Matches 70; Conservative 49; Mismatches 98; Indels 141; Gaps 19;

Qy 15 TVGEGPHWDHETQTLFYVDIVVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVS 74
Dy 17 TIG-----HVDHGKTKTLTAITMA-----LAARGGAKGR----- 45
Qy 75 LEREIALTLWDGVSAAPTSEIAIVNVE-PIKNNRLN-----DGKADPLGNLWTCM 125
Dy 46 -----KYDDISAFEEKARGITINTAHVEYETENRHYAHVDCPGHADYVKNMTCAA 97
Qy 126 AIDAGLPV-----GPVTGSLYHLGADKKVKMHESNIAIANGLAWSNCLKMYIYDSGKRR 180
Dy 98 QMDGGLVSGADGPMPTQREHLLAKQV-----GVPNIVVELN-----K 137
Qy 181 VDEYDYDASTLSISNQRPPLTFTEKHEVPGYPDGTIDEENGLWVAVFGQRIKIS--TQ 238
Dy 138 EDQVDY-DAELLELELEIRETLDRYEPFG-----DE-----IPIISGSALLALEALTE 184
Qy 239 QPEV-----LLDPVK--IPDPQ-----VT---SVAFG--- 260
Dy 185 NPEKPGDNKWKDYIYNLMQVDYSIPTPERETKPFIMAVEDVFSITGRGTATGRVER 244
Qy 261 -----GNPL-----DELHVTSLAGLQDSSLDKSLVNGHVRYRTGLGVKGFAGVKVK 307
Dy 245 GCVKIGTVELVGLRDKTKTTVTGLMFQKLTDESVAQDNV-----GILLRGVQKIDIE 298

RESULT 12
PGK_HALVA STANDARD; PRT; 401 AA.
AC P50315;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK.
OS Haloarcula vallismortis.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=28442;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29715;
RX MEDLINE=96197399; PubMed=8616244;
RA Brinkmann H., Martin W.;
RT "Higher-plant chloroplast and cytosolic 3-phosphoglycerate kinases: a case of endosymbiotic gene replacement.";
RL Plant Mol. Biol. 30:65-75(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate -> ADP + 3-phospho-D-glyceroyl phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBCELLULAR LOCATION: Cytosolic.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
```

```
CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L47295; AAB03731.1; -.
CC HSSP; P18912; IPHP.
CC InterPro; IPR001576; PGK.
CC Pfam; PF00162; PGK; 1.
CC PRINTS; PR00477; PHGLYCKINASE.
CC PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW Transferase; Kinase; Glycolysis.
SQ SEQUENCE 401 AA; 44193 MW; ACDD98F8856B490E CRC64;

Query Match 5.8%; Score 94; DB 1; Length 401;
Best Local Similarity 22.1%; Pred. No. 3.5;
Matches 78; Conservative 42; Mismatches 137; Indels 96; Gaps 18;

Qy 2 GP-VVEKIAELGKTV-----GEGPHWDHETQTLFYVDIVVEKTFHKYVPSQKKY 49
Dy 94 GPQAIHDIAIDLSGDVLVLENTMCDDELPEEDPEVKAQTEFVKTLAGEFDAYINDAYS 153
Qy 50 TFCKVDKLVSLIPLAGSPGRFVVSLERE-----TAILTWDG---VSAAPTSEIAIVNVEP 102
Dy 154 AHRSHASLVGFLPYMDAYAGR-VMETEYEANTAEKEFDGQVTVMVGGTKRATVIDVMT 212
Qy 103 HIKNNRLNDGKADP--LGNLWGTGMAIDAGLPVG-----PVTGSLYHLGADKKVKM--- 151
Dy 213 HL-----DEKVDDELGGI-AGTVPAAGHPGVGYDDANLYDEQWENSEKIESMLED 265
Qy 152 HESNIAIANGLAWSN-----DLKKMYIYDSGKRRVDEYDYDASTLSISNQRP 198
Dy 266 HRDQITLAVDLAYEDENDRAEQAVDDIDEKRLSYLDVSGSETLMEYS----- 312
Qy 199 LFTEKHEVPGYPDGTIDEENGLWVAVFGQRIKISTQPEVLLDTVKIPDPQVTSVA 258
Dy 313 -----PIIRESEAVFGEGR--AGMFEDEFSGTAGVLEAIADT-----DCFSW 355
Qy 259 FGGPNDELHVTSLAGLQDSSLDKSLVNGHVRYRTGLGVKGFAGVKVK 306
Dy 356 GGGD-----TSRAIEMYGNEDEF---GHVSIAGGAYIRALTRAQLVGVEV 398

RESULT 13
FAT2_RAT STANDARD; PRT; 4351 AA.
AC O8827;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin Fat 2 precursor (Multiple epidermal growth factor-like domains 1).
GN FAT2 OR MEGF1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
CC -1- SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB011527; BAA32458.1; ..
HSP; P08709; 18F9
InterPro; IPR002126; Cadherin.
InterPro; IPR000561; EGF-like.
InterPro; IPR000742; EGF_2.
InterPro; IPR001791; Laminin_G.
Pfam; PF00008; Cadherin; 33.
Pfam; PF00008; EGF; 2.
Pfam; PF00054; laminin_G; 1.
PRINTS; P00205; CADHERIN.
SMART; SM01112; CA; 33.
SMART; SM01181; EGF; 2.
SMART; SM0282; Lam; 1.
PROSITE; PS00232; CADHERIN_1; 14.
PROSITE; PS0268; CADHERIN_2; 33.
PROSITE; PS00225; LAM_G_DOMAIN; 1.
PROSITE; PS00222; EGF_1; 2.
PROSITE; PS01186; EGF_2; 2.
Transmembrane; Glycoprotein; Repeat; EGF-like domain; Signal.
SIGNAL 1 18
CHAIN 19 4351
DOMAIN 19 4050
TRANSMEM 4051 4071
DOMAIN 4072 4351
DOMAIN 34 148
DOMAIN 149 256
DOMAIN 363 458
DOMAIN 459 564
DOMAIN 565 669
DOMAIN 716 820
DOMAIN 821 925
DOMAIN 926 1032
DOMAIN 1033 1142
DOMAIN 1143 1242
DOMAIN 1243 1346
DOMAIN 1350 1448
DOMAIN 1449 1555
DOMAIN 1556 1660
DOMAIN 1661 1758
DOMAIN 1759 1872
DOMAIN 1873 1968
DOMAIN 1969 2070
DOMAIN 2071 2171
DOMAIN 2172 2272
DOMAIN 2273 2379
DOMAIN 2380 2481
DOMAIN 2482 2585
DOMAIN 2586 2692
DOMAIN 2693 2799
DOMAIN 2800 2908
DOMAIN 2909 3013
DOMAIN 3014 3115
DOMAIN 3116 3220
DOMAIN 3221 3323
DOMAIN 3324 3428
DOMAIN 3429 3533
DOMAIN 3534 3631
DOMAIN 3775 3946
DOMAIN 3952 3986
DOMAIN 3991 4024
DOMAIN 39 39
CARBOHYD 210 210
CARBOHYD 280 280
CARBOHYD 330 330
CARBOHYD 459 459
CARBOHYD 568 568
CARBOHYD 627 627
CARBOHYD 789 789
CARBOHYD 996 996

FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1276 1276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1417 1417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1899 1899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1998 1998 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2007 2007 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2102 2102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2165 2165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2183 2183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2325 2325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2368 2368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2387 2387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2430 2430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2470 2470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2547 2547 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2597 2597 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3127 3127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3278 3278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3312 3312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3432 3432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3603 3603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3770 3770 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3774 3774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3815 3815 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3842 3842 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3875 3875 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3906 3906 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3991 3991 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 4351 AA; 480644 MW; C564E7F9EBB3888 CRC64;
Query Match 5.7%; Score 92.5; DB 1; Length 4351;
Best Local Similarity 20.9%; Pred. No. let02;
Matches 76; Conservative 54; Mismatches 133; Indels 101; Gaps 17;
QY 17 GEGPHWDHETQTLTYFYVDVTVEKTFKXVPSQKKYTFCK-----v 54
Db 2667 GGPFWHDSLVPLVRLQVVPNEIPLPKF--SEPLTFPSAPEDLPGESEIGSVKVAADPPII 2724
QY 55 DKLVSEFTIPLAGSPGRFVSVSLEREIAL-----TW-----DGVSAAAP-----TSIEAI 97
Db 2725 YSLVQGTTPESNDD--VFSLDQDTGVLKVRKAMHSTKWKYQIDLMHACHIEDTDLVSL 2782
QY 98 VNVPEPHKNNRLNDGK-----ADPLGNLWGTMTAIDAGLPVGTGSLYHLGADKKV---- 149
Db 2783 VSVSIQVED--VNDNREVFADPYKAFLENN--PGTTVIQVTDANDQDTGSDGQSVYRL 2838
QY 150 -----KMHESTAIANGLAWSNDLKKM-----YYIDSGRRVDEYDYDASTLSIS 194
Db 2839 SVPEGSNIHEL-FAVDSESGWITTLQELDCETQOTVRYVVA-----FDH-GQTIQLS 2889
QY 195 NQRPLFTFEKHEVPGYPDGTIDEENLWVAVFGQRIIKISTQOPEVLDTVKIPDP-- 252
Db 2890 SQALVEVSIITDENDNPPRFASEDYRGSV-----VENNEPGEVLATLKTLDADV 2937
QY 253 -----QVTSVAFGGPNLDELHVTSGALQLDSS---LDKSLVNGHVVRVTGLGVKPGAV 304
Db 2938 SDQNRQVTCYITEGDPGLQFSISQVGDENRISRRKTLDRHIAKYLRLVTSADGKFOASV 2997
QY 305 KVKL 308
Db 2998 PVEV 3001
RESULT 14
FAT2_HUMAN
ID FAT2_HUMAN STANDARD; PRT; 4349 AA.
AC Q9NVQ8; O75091; Q9NSR7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin Fat 2 precursor (hFat2) (Multiple epidermal growth
DE factor-like domains 1).
GN FAT2 OR MEGF1 OR CDHF8.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20202599; PubMed=10716726;
RA Wu Q., Maniatis T.;
RT "Large exons encoding multiple ectodomains are a characteristic
FT feature of protocadherin genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
RN [2]
RP SEQUENCE OF 3777-4349 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple
FT EGF-like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
RN [3]
RP SEQUENCE OF 4142-4349 FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Wellenreuther R., Meyes H.-W., Weil B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 32 CADHERIN DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF231022; AAF61928.1; -;
DR EMBL; AB011535; BAA32463.1; -;
DR EMBL; AL157443; CAB75663.1; -;
DR Genew; HGNC:3596; FAT2.
DR MIN; 604269; -;
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001791; Laminin-G.
DR Pfam; PF00028; cadherin; 33.
DR Pfam; PF00008; EGF; 2.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 32.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00282; LamG; 1.
DR PROSITE; PS00232; CADHERIN_1; 14.
DR PROSITE; PS0268; CADHERIN_2; 32.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
KW Transmembrane; Glycoprotein; Repeat; EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 4349
FT DOMAIN 19 4048
FT TRANSMEM 4049 4069
FT DOMAIN 4070 4349
FT DOMAIN 34 148
FT DOMAIN 149 256
FT DOMAIN 363 458
FT DOMAIN 459 564
FT DOMAIN 565 669
FT DOMAIN 716 820
FT DOMAIN 821 925
FT DOMAIN 926 1032
FT DOMAIN 1033 1137
FT DOMAIN 1138 1242
FT DOMAIN 1243 1346
FT DOMAIN 1347 1448
FT DOMAIN 1449 1555
FT DOMAIN 1556 1660
FT DOMAIN 1661 1758
FT DOMAIN 1759 1872
FT DOMAIN 1873 2070
FT DOMAIN 2071 2171
FT DOMAIN 2172 2272
FT DOMAIN 2273 2379
FT DOMAIN 2380 2481
FT DOMAIN 2482 2585
FT DOMAIN 2586 2691
FT DOMAIN 2692 2797
FT DOMAIN 2798 2906
FT DOMAIN 2907 3011
FT DOMAIN 3012 3113
FT DOMAIN 3114 3218
FT DOMAIN 3219 3321
FT DOMAIN 3322 3426
FT DOMAIN 3427 3531
FT DOMAIN 3532 3642
FT DOMAIN 3643 3773
FT DOMAIN 3774 3944
FT DOMAIN 3945 3984
FT DOMAIN 3985 4022
FT DOMAIN 4023 4022
FT CARBOHYD 39 39
FT CARBOHYD 210 210
FT CARBOHYD 280 280
FT CARBOHYD 330 330
FT CARBOHYD 459 459
FT CARBOHYD 568 568
FT CARBOHYD 627 627
FT CARBOHYD 655 655
FT CARBOHYD 789 789
FT CARBOHYD 996 996
FT CARBOHYD 1175 1175
FT CARBOHYD 1303 1303
FT CARBOHYD 1383 1383
FT CARBOHYD 1417 1417
FT CARBOHYD 1904 1904
FT CARBOHYD 1998 1998
FT CARBOHYD 2007 2007
FT CARBOHYD 2165 2165
FT CARBOHYD 2183 2183
FT CARBOHYD 2325 2325
FT CARBOHYD 2368 2368
FT CARBOHYD 2387 2387
FT CARBOHYD 2430 2430
FT CARBOHYD 2470 2470
FT CARBOHYD 2547 2547
FT CARBOHYD 2597 2597
FT CARBOHYD 2654 2654
FT CARBOHYD 3125 3125
FT CARBOHYD 3276 3276
FT CARBOHYD 3310 3310
FT CARBOHYD 3430 3430
FT CARBOHYD 3471 3471
FT CARBOHYD 3601 3601
FT CARBOHYD 3772 3772
FT CARBOHYD 3813 3813
FT CARBOHYD 3840 3840
FT CARBOHYD 3873 3873
FT CARBOHYD 3904 3904
FT CARBOHYD 3989 3989
FT CARBOHYD 4117 4117
FT CONFLICT 4160 4160
SQ SEQUENCE 4349 AA; 479383 MW; 6C050CA17AD398ED CRC64;
Query Match 5.7%; Score 92; DB 1; Length 4349;
Best Local Similarity 21.9%; Pred. No. 1.le+02;
Matches 80; Conservative 41; Mismatches 140; Indels 104; Gaps 16;
QY 17 GEGPHWDHETOTLYFEVDVTEKTFHKYVPSQKKTFC-----V 54


```

Db 2666 GPPHNSLVVRLOQVPKKVSLPKF--SEPLYTFSAPEDLPEGSEIGIVKAAQDPVI 2723
QY 55 DKLVSFTIPLAGSGRFRVSLERIEAIL---TWGVSAAPTSIEAIVN-----V 100
Db 2724 YSLVRGTPESNKDG--VFSLDPTGVIKVRKPMDEHSTKLYQIDVMAHCLQNTDVVSLV 2781
QY 101 EPHKNNRLNDGK-----ADPIGNLWTGTMAIDAGLPVGPVT-----GSL-YHLGAD 146
Db 2782 SVNIQGVGDVNDNRVFEADPYKAVLTENMPV--GTSVIQVTAIDKDGTRDGVSYRLSAD 2839
QY 147 KVKMHESNIAIANGLAWSNDKKMYVIDSGKRVRDEYDVDASTLSISNQRPLTFEKKHE 206
Db 2840 PGSNVHEL-FAIDSESGWITLQEL-----DCETCQ-----TYHFH- 2874
QY 207 VPGYPDGGTI-----DEEGLMWAVFGQRIKISTQQPEVLDDTVKIPDP- 252
Db 2875 VVAYDHGQTIQLSSQALVQVSITDENDNAPRFASEYRGVSVENSEPGLVATLKTLDAD 2934
QY 253 -----QVTSVAFGPNLDELHVTISAGLQDDSS---LDKSLVNGHVRYRTGLGVKGFAG 303
Db 2935 ISEQRQVTCYITEGDPGLGFQISGVGDWRISRKTLDRHTAKYLLRLVTASDGKFOAS 2994
QY 304 VKVKL 308
Db 2995 VTVEI 2999

RESULT 15
RL1_STRVG STANDARD; PRT; 242 AA.
AC P48351;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L1.
GN RPLA.
OS Streptomyces virginiae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1961;
RN [1]
RP MEDLINE=96257210; PubMed=8675024;
RA Katayama M., Sakai Y., Okamoto S., Ihara F., Nihira T., Yamada Y.;
RT "Gene organization in the ada-rpl region of Streptomyces virginiae.";
RL Gene 171:135-136(1996).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC LOCATED IN THE NEIGHBORHOOD OF THE SITE WHERE ELONGATION FACTOR TU
CC IS BOUND TO THE RIBOSOME (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D50624; BAA09303.1; -
CC HSSP; P27150; 1AD2.
DR InterPro: IPR002143; Ribosomal_L1.
DR Pfam: PF00687; Ribosomal_L1; 1.
DR ProDom: PD001314; Ribosomal_L1; 1.
DR TIGRFAMs: TIGR01169; rplA_bact; 1.
DR PROSITE: PS01199; RIBOSOMAL_L1; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 242 AA; 26015 MW; 2194AGFF6FFB13CE CRC64;

Query Match 5.7%; Score 91.5; DB 1; Length 242;
Best Local Similarity 26.0%; Pred. No. 2.8;
Matches 41; Conservative 14; Mismatches 63; Indels 35; Gaps 5;
```

```

QY 25 ETOTLYFVDIVETFKHYVPSQKKYTFCKVDKLVSLFIPLAGSPGRFVVSLEREIALTW 84
Db 31 ETSTTTKFDGTVEVAFRLGVDPR-----KADQMVRGTVNLPHGTGK-----TARVLVFATG 80
QY 85 DGVSAAPTSTIEAIVNVEPHI-----KNNRLND-----CKADPLG-----N 119
Db 81 DRAAAAEAGADIVGDDELINEIAKGNRLNEFDVAVVATPDLMSKVGRIGRLGRLGPRCLMPN 140
QY 120 LWTGTMAIDAGLPVGPVPTGSLYHLGADKKVKMH 152
Db 141 PKTGTMTMDVAKAVTEIKGGKIEFRVDKHSNLH 173

Search completed: August 1, 2003, 14:32:45
Job time : 27 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:26:38 ; Search time 70 Seconds
(without alignments)
906.607 Million cell updates/sec

Title: US-10-089-986-2

Perfect score: 1615

Sequence: 1 MGPPVEKIAELGKTYTGEGP.....HVRVTGLGVKGFAGVKVL 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_virus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1615	100.0	308	5 Q95Y14	Q95Y14 photinus py
2	889.5	55.1	309	5 Q8TA67	Q8TA67 luciola cru
3	859.5	53.2	307	5 Q8TA68	Q8TA68 luciola lat
4	568	35.2	303	5 Q9NDS6	Q9NDS6 drosophila
5	568	35.2	303	5 Q9VFG5	Q9VFG5 drosophila
6	550	34.1	303	5 Q9VYR1	Q9VYR1 drosophila
7	538	33.3	306	5 Q9NDP1	Q9NDP1 sarcophaga
8	521	32.3	306	5 Q9BID6	Q9BID6 calliphora
9	474.5	29.4	299	13 Q91922	Q91922 xenopus lae
10	470.5	29.1	299	13 Q91BA8	Q91BA8 xenopus lae
11	455	28.2	299	13 Q91923	Q91923 gallus gall
12	453.5	28.1	299	11 Q925W3	Q925W3 rattus norv
13	445.5	27.6	267	5 Q95PD9	Q95PD9 calliphora
14	423.5	26.2	569	16 Q92ZR8	Q92ZR8 rhizobium m
15	409.5	25.4	584	16 Q8UOL2	Q8UOL2 agrobacteri
16	397.5	24.6	271	17 Q96Z19	Q96Z19 sulfolobus

17	365.5	22.6	285	17	Q97VC7	Q97VC7 sulfolobus
18	350.5	21.7	282	17	Q9HKF3	Q9HKF3 thermoplasm
19	350	21.7	285	17	Q96XG1	Q96XG1 sulfolobus
20	348	21.5	211	5	Q8T025	Q8T025 sarcophaga
21	339.5	21.0	286	16	Q9LOA7	Q9LOA7 streptomyce
22	331	20.5	289	16	Q9A9Z1	Q9A9Z1 caulobacter
23	315	19.5	293	16	Q9A7B9	Q9A7B9 caulobacter
24	308	19.1	282	17	Q97A53	Q97A53 thermoplasm
25	294.5	18.2	303	16	Q8XV54	Q8XV54 ralstonia s
26	292	18.1	285	17	Q97UH7	Q97UH7 sulfolobus
27	288.5	17.9	311	16	Q931B6	Q931B6 rhizobium m
28	287.5	17.8	296	16	Q8XRJ9	Q8XRJ9 ralstonia s
29	275.5	17.1	294	16	Q92T70	Q92T70 rhizobium m
30	273	16.9	322	16	Q8U8A8	Q8U8A8 agrobacteri
31	270	16.7	301	16	Q986F6	Q986F6 rhizobium l
32	256	15.9	314	16	Q8UHT0	Q8UHT0 agrobacteri
33	228.5	14.1	294	16	Q92RN9	Q92RN9 rhizobium m
34	213	13.2	302	16	Q82MM6	Q82MM6 salmonella
35	206	12.8	305	16	Q8U8X6	Q8U8X6 agrobacteri
36	203	12.6	302	16	Q8Z4I8	Q8Z4I8 salmonella
37	202.5	12.5	291	16	Q8YE81	Q8YE81 brucella me
38	201	12.4	290	16	Q98DE0	Q98DE0 rhizobium l
39	153.5	9.5	348	16	Q8U8E1	Q8U8E1 agrobacteri
40	145.5	9.0	296	16	Q984W1	Q984W1 rhizobium l
41	140	8.7	303	16	Q926B8	Q926B8 rhizobium m
42	132	8.2	304	16	Q926I2	Q926I2 rhizobium m
43	126.5	7.8	115	11	Q9QWP2	Q9QWP2 rattus norv
44	124	7.7	353	16	Q8XRK7	Q8XRK7 ralstonia s
45	122.5	7.6	306	16	Q98EA9	Q98EA9 rhizobium l

ALIGNMENTS

RESULT 1

Q95Y14
ID Q95Y14 PRELIMINARY; PRT; 308 AA.
AC Q95Y14;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Luciferin regenerating enzyme.
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Lampyrinae; Photinus.
OX NCBI_TaxID=7054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21453302; PubMed=11457857;
RA Gomi K., Kajiyama N.;
RT "Oxyluciferin, a Luminescence Product of Firefly Luciferase, Is
Enzymatically Regenerated into Luciferin.";
RL J. Biol. Chem. 276:36508-36513(2001).
DR EMBL; AB062786; BAB60700.1; -;
SQ SEQUENCE 308 AA; 33806 MW; 4403A6EC6459A378 CRC64;

Query Match	100.0%;	Score 1615;	DB 5;	Length 308;
Best Local Similarity	100.0%;	Pred. No. 1.8e-125;	Mismatches 0;	Indels 0;
Matches 308;	Conservative 0;			
QY	1	MGPPVEKIAELGKTYTGEGP	HDHETQTLVFVDTVEKTFHKYVP	SQKKTFCCKVDKLYSF 60
Db	1	MGPPVEKIAELGKTYTGEGP	HDHETQTLVFVDTVEKTFHKYVP	SQKKTFCCKVDKLYSF 60
QY	61	IIPLAGSGPRFVSVLEREIAL	TWDGVSAAPTISIEAIVNVEPH	IKNNRNDKADPLGNL 120
Db	61	IIPLAGSGPRFVSVLEREIAL	TWDGVSAAPTISIEAIVNVEPH	IKNNRNDKADPLGNL 120
QY	121	WTGTTAIDAGLPVPGVTGSL	YHLGADKKVKKMHESNIA	TLANGLAWSNDLKKYYIDSGRR 180
Db	121	WTGTTAIDAGLPVPGVTGSL	YHLGADKKVKKMHESNIA	TLANGLAWSNDLKKYYIDSGRR 180

QY 181 VDEYDYDASTLSISNORPLTFEKEHVPYDQGTIDEENLWVAVFOGRIIKISTQOP 240
 DB 181 VDEYDYDASTLSISNORPLTFEKEHVPYDQGTIDEENLWVAVFOGRIIKISTQOP 240
 QY 241 EVLDDTVKIPDQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLGVKG 300
 DB 241 EVLDDTVKIPDQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLGVKG 300
 QY 301 FAGVKVKL 308
 DB 301 FAGVKVKL 308

RESULT 2

Q8TA67 PRELIMINARY; PRT; 309 AA.
 AC Q8TA67;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Luciferin-regenerating enzyme.
 G-LRE.
 OS Luciola cruciata (Japanese firefly) (Genji firefly).
 GN Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
 NCBI_TaxID=7051;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gomi K., Hirokawa K., Kajiyama N.;
 RT "Molecular cloning and expression of luciferin-regenerating enzyme
 (LRE) cDNA in Luciola cruciata and Luciola lateralis.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB072448; BAB85479.1; -
 SQ SEQUENCE 309 AA; 33604 MW; BE33C4E21EF51E6B CRC64;

Query Match 55.1%; Score 889.5; DB 5; Length 309;
 Best Local Similarity 56.6%; Pred. No. 1.8e-65;
 Matches 176; Conservative 47; Mismatches 83; Indels 5; Gaps 4;

QY 1 MGPVVEKIAELGKYTVGEGPHMDHETQTLFYFVDTVEKTFHKYVPSQKKTFCVKDLVSF 60
 DB 1 MAPTVEQIVELGTLVLAESPWHDDTOSLYFVDVGRSVNKYVPTTHTQLKFDKPNPSF 60
 QY 61 IIPLAGSGRFVVSLEIRIATLWDGVSAAPTSIE--AIVNVEPHIKNNRLNDGKADPLG 118
 DB 61 IIPVKGCSDFIVSLERINLLTWDGASSAPSKIEKIAVFNTPKSENRLNDGKADPLG 120
 QY 119 NLWCTMAIDAGLPVG-PVTSGLYHLGADKKVKMHESNIAIANGSLNLDKMKYYIDSG 177
 DB 121 NLWAGTMNNGSDHTTGTTPVRGTLSSL-SNKQVKEHVSVCISNGLAWSKDLKREYIDSA 179
 QY 178 KRVDEYDYDASTLSISNORPLTFEKEHVPYDQGTIDEENLWVAVFOGRIIKIST 237
 DB 180 VRQVDQDFDANKLSLRQPLTFDKHGMSPDQGTIDEENLWATCGDKVLKIDT 239
 QY 238 QQPEVLLDTVKIPDQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLG 297
 DB 240 STPEILLGIVEIPEHQVTSVCIGGAELNVLVYTTASIKLPAGDETKPM-KGAIYKVTGLG 298
 QY 298 VKGFGVKVKL 308
 DB 299 VKGLPGDRVKL 309

RESULT 3

Q8TA68 PRELIMINARY; PRT; 307 AA.
 AC Q8TA68;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Luciferin-regenerating enzyme.

GN H-LRE.
 OS Luciola lateralis (Firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
 NCBI_TaxID=7052;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gomi K., Hirokawa K., Kajiyama N.;
 RT "Molecular cloning and expression of luciferin-regenerating enzyme
 (LRE) cDNA in Luciola cruciata and Luciola lateralis.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB072447; BAB85478.1; -
 SQ SEQUENCE 307 AA; 34286 MW; B03CED0AC609A92C CRC64;

Query Match 53.2%; Score 859.5; DB 5; Length 307;
 Best Local Similarity 52.8%; Pred. No. 5.5e-63;
 Matches 163; Conservative 58; Mismatches 85; Indels 3; Gaps 3;

QY 1 MGPVVEKIAELGKYTVGEGPHMDHETQTLFYFVDTVEKTFHKYVPSQKKTFCVKDLVSF 60
 DB 1 MSPVIEQITEVDNFOIGEGPHMDTETQSLYFVDILEKSIHKYVPSKQHTKMLNKRPSF 60
 QY 61 IIPLAGSGRFVVSLEIRIATLWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLG 120
 DB 61 IIPKETSDFEIVSLERDLCVLTWDGVSATPSHLETIVTDTCIEGNTFNDGKADAFG 120
 QY 121 WTGTMAIDAGL-PVGPVTSGLYHLGADKKVKMHESNIAIANGSLNLDKMKYYIDSGKR 179
 DB 121 WAGTLYSKFDIEKQGNPTGTLVSL-SNKQLRKHISNIFLSNGLANKDSKKEFFIDSNKR 179
 QY 180 RVDEYDYDASTLSISNORPLTFEKEHVPYDQGTIDEENLWVAVFOGRIIKISTQO 239
 DB 180 TIDQFDYDSENLIIISNCQPLFTLDKHGICQPLDPAQTIDENDNLWVAIVRGKVINIGTKO 239
 QY 240 PEVLDTVKIPDQVTSVAFGPNLDELHVTSLAGLQDSSLDKSLVNGHVYRTGLGVK 299
 DB 240 PESSLLGVINPESLITSVCGSKLDLIVTTSGIKEYETDSTK-LVKGGLRVTVGLGVK 298
 QY 300 GFGVKVKL 308
 DB 299 GLPAHRFSL 307

RESULT 4

Q9NDS6 PRELIMINARY; PRT; 303 AA.
 ID Q9NDS6
 AC Q9NDS6
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Senescence marker protein-30 (SMP-30).
 GN SMP-30 OR CG7390.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON S; TISSUE-WHOLE BODY;
 RX PubMed-10817937;
 RA Goto S.G.;
 RT "Expression of Drosophila homologue of senescence marker protein-30
 during cold acclimation.";
 RL J. Insect Physiol. 46:1111-1120(2000).
 DR EMBL; AB029490; BAA92938.1; -
 DR FlyBase; FBgn038257; smg-30.
 SQ SEQUENCE 303 AA; 33349 MW; 68F12B92E1AAB5AD CRC64;

Query Match 35.2%; Score 568; DB 5; Length 303;
 Best Local Similarity 40.7%; Pred. No. 6.9e-39;
 Matches 121; Conservative 48; Mismatches 116; Indels 12; Gaps 5;

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Clapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AE003706; AAF55095.1; -;
DR EMBL; AY071130; AA48752.1; -;
DR FlyBase; FBgn0038257; smg-30.
SQ SEQUENCE 303 AA; 33377 MW; 68F92692E1AC88AD CRC64;

Best Local Similarity	40.7%;	Pred. NO.	6.9e-39;
Matches	121;	Conservative	48;
Mismatches	116;	Indels	12;
Gaps	5		

Qy	16	VGEGPHDHDHETQTLFYVDVTVEKTHKHYPSOKKTYFCVKDK--VVSFFIIPLAGSGPGEVW 73
Db	15	LGEGPHDWDVDSQSLYYVLDSESAGINRYDFKQNKYVRAKIEGFASFATPLPVENKPKOEFAV 74
Qy	74	SLEREIAILTWGDSYSAAPTSTEATVNVNPHLKNRNLNDGKADPCLGNLTMTGTMAIDAGLPV 133
Db	75	CGGLRTVIVQWDGVSAYAKVPTTFEEVQPDLEKNRLNDAKTDPNGRFFYGGTMA--DSGDIF 133
Qy	134	GPVTGSLYHLGADKKVKMHESNIAIANGLANSDNLKKWYYIDSGKRRRVDEYDSDASTLJI 193
Db	134	TQWKGELYSWOAGGQPNAIRSKVGTSGNLANDVRAKKEYFIDTNNHEVLAYDYNOSTGAV 193
Qy	194	SNORPLPTFEK--HEVPGYPDQOTIDEGNLWAVFQGORIIKITQOPEVLLDVTVPKIDP 251
Db	194	SNPKVIFDLRKIRPEGPLFDPMVDDGNIVVATFNNGTVFKVNPSTGKKILLE-IKIPT 252
Qy	252	PQVTSVAFGPGPNLDLHVTSLAGLQDSSLDKSLVNGHVYRVYGLGVKGFAGVKVKL 308
Db	253	TQITSVAFGPNLDILYVTTA-----NKFQPKPAGTTFQVTLGNAGYACGNLKI 303

RESULT	6
Q9VYR1	
ID	Q9VYR1
AC	PRELIMINARY; PRT; 303 AA.
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE	CG1803 protein (Regucalcin HOMOLOGUE).
GN	REGUCALCIN OR CG1803.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;

SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R.C., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Eppanckoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.-P., Bhendari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

Query Match 32.3%; Score 521; DB 5; Length 306;
Best Local Similarity 39.0%; Pred. No. 5.4e-35;
Matches 122; Conservative 44; Mismatches 127; Indels 20; Gaps 7;
QY 3 PVVEKTAELGKYTVGEGPHDHEOTLYFVDVTEKTFHKYVPSQKKYTFCKV--DKLVSF 60
DB 7 PLPDSRAEL-----GEGPHDVAATONDIYVDIINAGKLLRYNENKVVYKAKIGEDLAGF 61
QY 61 IIPLAGSPGRVSVSLEREIAILTWGVSAAPTSTIEAIVNV--EPHFKNNRLNDGKADPL 117
DB 62 AVPEVGTDDQFVSGRRVYVQDGVSETAKVLKTLTFEVTGDERETGNRFNDGKCDPR 121
QY 118 GNLTGTMAIDAGLPGVPTGSLYHLGADKKVKMHESNIAIANGLANSNDLKMYIYDSG 177
DB 122 GRIFACTMRY-VGDEFHRYGELYKYENGGEVEIKSDVGSINGLAWNERTKRFYIDTT 180
QY 178 KRRVDEYDYDASTLSISNORPLTFEKEHVPVGY--PDGQITDEGNLWAVFQGORLIKI 235
DB 181 DYEVEYDYDFETGKASPKVFNLRKTSFKDHLPLPDGMTIDTEGNDIYVATFNHSHIYKI 240
QY 236 STQOPEVLLTVKIPDPQVTSVAFGGPNLDELHVTSLAGLQDDSSLDKSLVNGHVYRTVG 295
DB 241 NPPTGQVLL-IFPKCKQITSAFAGGPNHFLFVTTT-----SRFCEPHPGAGTYYKVTG 293
QY 296 LGVKGFGAGVKVL 308
DB 294 LGAKGYPMTKIQL 306

RESULT 9

QY 91922 PRELIMINARY; PRT; 299 AA.
AC Q91922;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Regucalcin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20351777; PubMed=10891565;
RA Misawa H., Yamaguchi M.;
RT "The gene of Ca2+-binding protein regucalcin is highly conserved in
vertebrate species."
RL Int. J. Mol. Med. 6:191-196(2000).
DR EMBL: AB037936; BAA90694.1.
SQ SEQUENCE 299 AA; 33087 MW; BABB9749258BB2C7 CRC64;
Query Match 29.4%; Score 474.5; DB 13; Length 299;
Best Local Similarity 37.2%; Pred. No. 3.7e-31;
Matches 110; Conservative 46; Mismatches 117; Indels 23; Gaps 7;
QY 14 YTVGEGPHDHEOTLYFVDVTEKTFHKYVPSQKKYTFCKVKLVFIPLAGSPGRFV 73
DB 14 YKIGESPVWEKEGTLFVDITGQKVCWRDPSTKKVQSVSVAPIGSV--ALRKS GGIVL 71
QY 74 SLEREIAILTWGVSAAPTSTIEAIVNVPEPHKNNRLNDGKADPLGNLTGTMTAIDAGLPV 133
DB 72 AMGNFESALNWE-----DQSVTTLARVDEDKPNNRFDKGVDPGREFLAGTMSQE---I 122
QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLANSNDLKMYIYDSGRRRVEYDYDA 189
DB 123 RPVAVERNQGSFLTLYPDHSHVHFMDVDSINGLDWSLDHKTLYIIDLKFKVDALDYDM 182
QY 189 STLSISNORPLTFEKEHVPVPGQITDEGNLWAVFQGORIHKISTQOPEV--LLDT 246
DB 183 KTGKSSNRRTLYKLOQDE--GIPDGMCIDAEGLWVACVNGRVRIRI---DPETGKQIQT 237
QY 247 VKIPDPQVTSVAFGGPNLDELHVTSLAGLQDDSSLDKSLVNGHVYRTVGLGVKGA 302
DB 238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEWKKRQPSGGGIYKISGLGVKGA 293

Query Match 29.4%; Score 474.5; DB 13; Length 299;
Best Local Similarity 37.2%; Pred. No. 3.7e-31;
Matches 110; Conservative 46; Mismatches 117; Indels 23; Gaps 7;

QY 14 YTVGEGPHDHEOTLYFVDVTEKTFHKYVPSQKKYTFCKVKLVFIPLAGSPGRFV 73
DB 14 YKIGESPVWEKEGTLFVDITGQKVCWRDPSTKKVQSVSVAPIGSV--ALRKS GGIVL 71
QY 74 SLEREIAILTWGVSAAPTSTIEAIVNVPEPHKNNRLNDGKADPLGNLTGTMTAIDAGLPV 133
DB 72 AMGNFESALNWE-----DQSVTTLARVDEDKPNNRFDKGVDPGREFLAGTMSQE---I 122
QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLANSNDLKMYIYDSGRRRVEYDYDA 189
DB 123 RPVAVERNQGSFLTLYPDHSHVHFMDVDSINGLDWSLDHKTLYIIDLKFKVDALDYDM 182
QY 189 STLSISNORPLTFEKEHVPVPGQITDEGNLWAVFQGORIHKISTQOPEV--LLDT 246
DB 183 KTGKSSNRRTLYKLOQDE--GIPDGMCIDAEGLWVACVNGRVRIRI---DPETGKQIQT 237

QY 247 VKIPDPQVTSVAFGGPNLDELHVTSLAGLQDDSSLDKSLVNGHVYRTVGLGVKGA 302
DB 238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEWKKRQPSGGGIYKISGLGVKGA 293
RESULT 10
QY 91922 PRELIMINARY; PRT; 299 AA.
AC Q91922;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Senescence marker protein-30.
GN XMP-30.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ANIMAL CAP;
RX MEDLINE=20193502; PubMed=10727865;
RA Sato A., Asashima M., Yokota T., Nishinakamura R.;
RT "Cloning and expression pattern of a Xenopus pronephros-specific gene,
XMP-30."
RL Mech. Dev. 92:273-275(2000).
DR EMBL: AB033368; BAA93719.1.
SQ SEQUENCE 299 AA; 33073 MW; BABB9749258BB41C CRC64;

Query Match 29.1%; Score 470.5; DB 13; Length 299;
Best Local Similarity 36.8%; Pred. No. 7.9e-31;
Matches 109; Conservative 47; Mismatches 117; Indels 23; Gaps 7;
QY 14 YTVGEGPHDHEOTLYFVDVTEKTFHKYVPSQKKYTFCKVKLVFIPLAGSPGRFV 73
DB 14 YKIGESPVWEKEGTLFVDITGQKVCWRDPSTKKVQSVSVAPIGSV--ALRKS GGIVL 71
QY 74 SLEREIAILTWGVSAAPTSTIEAIVNVPEPHKNNRLNDGKADPLGNLTGTMTAIDAGLPV 133
DB 72 AMGNFESALNWE-----DQSVTTLARVDEDKPNNRFDKGVDPGREFLAGTMSQE---I 122
QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLANSNDLKMYIYDSGRRRVEYDYDA 188
DB 123 RPVAVERNQGSFLTLYPDHSHVHFMDVDSINGLDWSLDHKTLYIIDLKFKVDALDYDM 182
QY 189 STLSISNORPLTFEKEHVPVPGQITDEGNLWAVFQGORIHKISTQOPEV--LLDT 246
DB 183 KTGKSSNRRTLYKLOQDE--GIPDGMCIDAEGLWVACVNGRVRIRI---DPETGKQIQT 237
QY 247 VKIPDPQVTSVAFGGPNLDELHVTSLAGLQDDSSLDKSLVNGHVYRTVGLGVKGA 302
DB 238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEWKKRQPSGGGIYKISGLGVKGA 293

Query Match 29.1%; Score 470.5; DB 13; Length 299;
Best Local Similarity 36.8%; Pred. No. 7.9e-31;
Matches 109; Conservative 47; Mismatches 117; Indels 23; Gaps 7;

QY 14 YTVGEGPHDHEOTLYFVDVTEKTFHKYVPSQKKYTFCKVKLVFIPLAGSPGRFV 73
DB 14 YKIGESPVWEKEGTLFVDITGQKVCWRDPSTKKVQSVSVAPIGSV--ALRKS GGIVL 71
QY 74 SLEREIAILTWGVSAAPTSTIEAIVNVPEPHKNNRLNDGKADPLGNLTGTMTAIDAGLPV 133
DB 72 AMGNFESALNWE-----DQSVTTLARVDEDKPNNRFDKGVDPGREFLAGTMSQE---I 122
QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLANSNDLKMYIYDSGRRRVEYDYDA 188
DB 123 RPVAVERNQGSFLTLYPDHSHVHFMDVDSINGLDWSLDHKTLYIIDLKFKVDALDYDM 182
QY 189 STLSISNORPLTFEKEHVPVPGQITDEGNLWAVFQGORIHKISTQOPEV--LLDT 246
DB 183 KTGKSSNRRTLYKLOQDE--GIPDGMCIDAEGLWVACVNGRVRIRI---DPETGKQIQT 237
QY 247 VKIPDPQVTSVAFGGPNLDELHVTSLAGLQDDSSLDKSLVNGHVYRTVGLGVKGA 302
DB 238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEWKKRQPSGGGIYKISGLGVKGA 293

RESULT 11

QY 91923 PRELIMINARY; PRT; 299 AA.
AC Q91923;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Regucalcin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20351777; PubMed=10891565;
RA Misawa H., Yamaguchi M.;

RT *the gene of Ca2+-binding protein regucalcin is highly conserved in
 RT vertebrate species";
 RL Int. J. Mol. Med. 6:191-196(2000).
 DR EMBL; AB037935; BAA90693.1; -;
 SQ SEQUENCE 299 AA; 33229 MW; 4754C7571164720E CRC64;

Query Match 28.2%; Score 455; DB 13; Length 299;
 Best Local Similarity 37.0%; Pred. No. 1.5e-29;
 Matches 112; Conservative 40; Mismatches 123; Indels 28; Gaps 8;

QY 13 KYTVGEGPHWDHETQTLFYVDVTEKTFHKYVPSQKYYTFCKVDKLVFIPIPLAGSPGRFV 72
 DB 13 RYRLGSPWDEKENSLLCDVITGRKVCWRDAAAGQVQALSDAPVSSV--ALRKSGDYV 70
 QY 73 VSLRETAITLWDGVSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDAGLP 132
 DB 71 ITLGTREFAALKW-----KEQLVTTIAQVDRDKANNRDNFGKVPAGRYFAGTMAEE---- 121
 QY 133 VGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIYDSGKRRVDEYDYD 187
 DB 122 IRPAVLRRQGSLLYTLCPDHSVVKHFQDQVDSINGLWSLQHKTFYFIDSLSYSDAFDYD 181
 QY 188 ASTLSISNORPLFTFEKHEVPGYDPGQTIDEEGNLWVAVFGQRIIKISTQQPEV--LLD 245
 DB 182 LQTKIGNRRSVYKLEREE--SIPDGMCIDTEGKLWVACDGGRVIRL---DPETGKRIQ 236
 QY 246 TVKLPDPQVTSVARGGNLDELHVTSLAGLQDDSSLDKSLVNGHVYRVTVGLGVKG----- 300
 DB 237 TVKLPVDPKTTSCCFGGKDYSEMYVTSASDGDNRWLSRQPOAGGVFKITGLGVKGIPYP 296
 QY 301 FAG 303
 DB 297 FAG 299

RESULT 12

Q925W3
 ID Q925W3 PRELIMINARY; PRT; 299 AA.
 AC Q925W3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Regucalcin.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Misawa H., Yamaguchi M.;
 RT "The gene family encoding the calcium-binding protein regucalcin";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB037934; BAA90692.1; -;
 SQ SEQUENCE 299 AA; 33390 MW; 508F2D95FCA4EE35 CRC64;

Query Match 28.1%; Score 453.5; DB 11; Length 299;
 Best Local Similarity 37.5%; Pred. No. 2e-29;
 Matches 111; Conservative 40; Mismatches 122; Indels 23; Gaps 7;

QY 14 YTVGEGPHWDHETQTLFYVDVTEKTFHKYVPSQKYYTFCKVDKLVFIPIPLAGSPGRFV 73
 DB 14 YRCGSPWNEEASKCLLEVDIPSKTVCRWDSISNRVQRVGDAPVSSV--ALRQSGGYVA 71
 QY 74 SLERETAITLWDGVSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLTGTMAIDAGLPV 133
 DB 72 TIGTFCALNWE-----DQSVFILAWDEKNNRDNFGKVDYDPAGRYFAGTMAEE---T 122
 QY 134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIYDSGKRRVDEYDYDA 188
 DB 123 APAVLERHOGSLYFLPDHRSVKKYFDQVDISNGLWSLDHKIFYYIDSLSYTVDADFYL 182
 QY 189 STLSTISNORPLFTFEKHEVPGYDPGQTIDEEGNLWVAVFGQRIIKISTQQPEV--LLD 246

Db 183 PTGQISNRRRTVYKMEKDE--QIPDGMCIDVEGKLWVACNGGRVIRL---DPETGKRLQT 237
 QY 247 VKLPDPQVTSVARGGNLDELHVTSLAGLQDDSSLDKSLVNGHVYRVTVGLGVKGFA 302
 DB 238 VKLPVDPKTTSCCFGGKDYSEMYVTCAROGMSAEGLLRQDPDAGNIFKAITGLGVKGIA 293

RESULT 13

Q93PD9
 ID Q93PD9 PRELIMINARY; PRT; 267 AA.
 AC Q93PD9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Anterior fat body protein (fragment).
 GN AFP.
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Calliphoridae; Calliphora.
 OX NCBI_TaxID=7373;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FAT BODY;
 RA Hansen I.A., Meyer S.R., Berlinger M.J., Scheller K.;
 RT "Identification of new interactors of the hexamerin receptor in the
 RT blowfly Calliphora vicina";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF326959; AAK51353.1; -;
 FT NON_TER 1
 SQ SEQUENCE 267 AA; 29836 MW; 76F4EEA957D032FF CRC64;

Query Match 27.6%; Score 445.5; DB 5; Length 267;
 Best Local Similarity 39.4%; Pred. No. 7.8e-29;
 Matches 102; Conservative 37; Mismatches 107; Indels 13; Gaps 5;

QY 55 DKLVFTIPLAGSPGRFVSVLSEREIAITLWDGVSAAPTSTEAIVNV---EPHINKNRLND 111
 DB 17 EDLAGFAVPVEGTTDQFVAGSGRRVIVQWDGVSETAQVLTFLFEVQTGDERFTGNRFND 76
 QY 112 GKADPLGNLTGTMAIDAGLPVGTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKM 171
 DB 77 GKCDPGRGLFAGTMRY--VGDEFHRHGYELKYENGGEVEVVKSDVGSINGLAWEKTKKF 135
 QY 172 YYIDSKRRRVDEYDYDASTLSISNORPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFG 229
 DB 136 YYIDTDDYEVKEYDYDFETCKASNPVFNLRKTSPPKHLLPDGMTIDTEGNYLVATFNG 195
 QY 230 QRIKISTQQPEVLLDTVTKIPDPQVTSVARGGNLDELHVTSLAGLQDDSSLDKSLVNGH 289
 DB 196 HSIYKINPTTQGVILLE-IKFPCKQITSAAFGGPNHILFVTTT-----SRFGEPHPAGT 248
 QY 290 VYRVTVGLGVKGFAVGVKVL 308
 DB 249 TYKVTGLGAGKGYPMTKIQL 267

RESULT 14

Q92ZR8
 ID Q92ZR8 PRELIMINARY; PRT; 569 AA.
 AC Q92ZR8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative calcium binding transcriptional regulatory protein.
 GN RA0381 OR SMA0717.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RC STRAIN=1021;
RA MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federpspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL: AE007229; AAK65039.1; -
DR InterPro: IPR00285; HPH_Ic1r.
DR Pfam: PF01614; Ic1r; 1.
DR Plasmid: Hypothetical protein; Complete proteome.
SQ SEQUENCE 569 AA; 60943 MW; EB34C90D654EE8AD CRC64;

Query Match 26.2%; Score 423.5; DB 16; Length 569;
Best Local Similarity 34.6%; Pred. No. 1.6e-26;
Matches 100; Conservative 46; Mismatches 102; Indels 41; Gaps 7;

Qy 16 VEGPWHDETHQTLFYVDVTEKTFHKYVPSQKTYTFCVKDKLVSVFIIPLAGSPGRFVVSL 75
Db 16 VEGPWHDETHQTLFYVDVTEKTFHKYVPSQKTYTFCVKDKLVSVFIIPLAGSPGRFVVSL 75
Qy 293 LGEAPVWVEREKLYWVDILHPAVHREDPVVTGKNESCNVAKLVSAVLP-----TR 342
Db 293 LGEAPVWVEREKLYWVDILHPAVHREDPVVTGKNESCNVAKLVSAVLP-----TR 342
Qy 76 ERETAILTWDCVS-----AAPTSTIEALVWVEPHIKNNRLNDGKADPLGNLWTG 123
Db 76 ERETAILTWDCVS-----AAPTSTIEALVWVEPHIKNNRLNDGKADPLGNLWTG 123
Qy 343 NEGILVASQDGVHEHFDGRGDNFPAEP-----EPGLPENRLNDKADVPDGLWVG 393
Db 343 NEGILVASQDGVHEHFDGRGDNFPAEP-----EPGLPENRLNDKADVPDGLWVG 393
Qy 124 TMAIDAGLPVGPVTHLGLADKKVKMHESNIAIANGLAWSNDLKKMYIDSKRRVDE 183
Db 124 TMAIDAGLPVGPVTHLGLADKKVKMHESNIAIANGLAWSNDLKKMYIDSKRRVDE 183
Qy 394 SMRLDVSPP-----TGSILRUTSAGEVTRAGFTVANGLAWSDDSTFYFVDVTPGIIYA 449
Db 394 SMRLDVSPP-----TGSILRUTSAGEVTRAGFTVANGLAWSDDSTFYFVDVTPGIIYA 449
Qy 184 YDYDASTLSISNORPLETFEKEHVPYDGPDTIDEENLWVAVFQGORIIRKISTQQPEVL 243
Db 184 YDYDASTLSISNORPLETFEKEHVPYDGPDTIDEENLWVAVFQGORIIRKISTQQPEVL 243
Qy 450 YDEAREGSIANRRVETVP--EAEGRPDGLAVDAGGVWCAIWDGWR---VNYRDPGR 504
Db 450 YDEAREGSIANRRVETVP--EAEGRPDGLAVDAGGVWCAIWDGWR---VNYRDPGR 504
Qy 244 LD-TVKIPDQVTSVAFGGPNLDLHVTSAGLQDLDSSLDKSLVNGHYR 291
Db 244 LD-TVKIPDQVTSVAFGGPNLDLHVTSAGLQDLDSSLDKSLVNGHYR 291
Qy 505 LDRAVELPVPRTSVAFGGDELATLFTSARTRLPASTLTLEAPLSGGIF 553
Db 505 LDRAVELPVPRTSVAFGGDELATLFTSARTRLPASTLTLEAPLSGGIF 553

RESULT 15
Q8UJL2 PRELIMINARY; PRT; 584 AA.
AC Q8UJL2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transcriptional regulator, IclR family/regucalcin.
GN ATU5464 OR AGR_PAT565.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.F., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmieu K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL: AE008967; AAL46151.1; ALT_INIT.
DR EMBL: AE007915; AAK90839.1; -
RW Plasmid: Complete proteome.
SQ SEQUENCE 584 AA; 62034 MW; 7E9F274048791037 CRC64;

Query Match 25.4%; Score 409.5; DB 16; Length 584;
Best Local Similarity 33.1%; Pred. No. 2.4e-25;
Matches 99; Conservative 57; Mismatches 110; Indels 33; Gaps 9;

Qy 16 VEGPWHDETHQTLFYVDVTEKTFHKYVPSQKTYTFCVKDKLVSVFIIPLAGSPGRFVVSL 75
Db 16 VEGPWHDETHQTLFYVDVTEKTFHKYVPSQKTYTFCVKDKLVSVFIIPLAGSPGRFVVSL 75
Qy 308 LGEAPVWVEREKLYWVDILHPAVHREDPVVTGKNESCNVAKLVSAVLP-----GRLLVAS 365
Db 308 LGEAPVWVEREKLYWVDILHPAVHREDPVVTGKNESCNVAKLVSAVLP-----GRLLVAS 365
Qy 76 ERETAILTWDCV---SPAPTSIEALVWVEPHIKNNRLNDGKADPLGNLWTGTMADGLP 132
Db 76 ERETAILTWDCV---SPAPTSIEALVWVEPHIKNNRLNDGKADPLGNLWTGTMADGLP 132
Qy 366 Q-----DGVWELDFASGRLTTPFVSPEAGIADNRLNDKACGPDGAIWVGSMDASKP 417
Db 366 Q-----DGVWELDFASGRLTTPFVSPEAGIADNRLNDKACGPDGAIWVGSMDASKP 417
Qy 133 VGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSKRRVDEYDASTLS 192
Db 133 VGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSKRRVDEYDASTLS 192
Qy 418 ----TGALYRINANGASERKEGGIIVSNGLWSPDGRTFYFVDVTPGLIHAYDCDPAFGA 473
Db 418 ----TGALYRINANGASERKEGGIIVSNGLWSPDGRTFYFVDVTPGLIHAYDCDPAFGA 473
Qy 193 ISNORPLETFEKEHVP---GYPDGQTIIDEENLWVAVFQGORIIRKISTQQPEVLDD-TVK 248
Db 193 ISNORPLETFEKEHVP---GYPDGQTIIDEENLWVAVFQGORIIRKISTQQPEVLDD-TVK 248
Qy 474 LSQRR-----EFARIPVADGRPDGLAVDAEGGVWCAIWDGWCVR---YLPNGKLDQVID 525
Db 474 LSQRR-----EFARIPVADGRPDGLAVDAEGGVWCAIWDGWCVR---YLPNGKLDQVID 525
Qy 249 IPDQVTSVAFGGPNLDLHVTSAGLQDLDSSLDKSLVNGHYR-----VTGLGVKGFAG 303
Db 249 IPDQVTSVAFGGPNLDLHVTSAGLQDLDSSLDKSLVNGHYR-----VTGLGVKGFAG 303
Qy 526 MPVPRPSSIAFGPDLSTLFTSARTRLPASTLTADAPLSGLGFCRPGISGARIAMFEG 584
Db 526 MPVPRPSSIAFGPDLSTLFTSARTRLPASTLTADAPLSGLGFCRPGISGARIAMFEG 584
```

Search completed: August 1, 2003, 14:34:10
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:32:13 ; Search time 24 Seconds
(without alignments)
377.594 Million cell updates/sec

Title: US-10-089-986-2

Perfect score: 1615

Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVVRYTGLGVKGFGAGVKVL 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	6.8	332	4	US-09-134-001C-4646
2	94	5.8	413	4	US-09-134-001C-5111
3	93	5.8	396	4	US-08-501-126-2
4	91	5.6	397	2	US-08-282-197C-55
5	90	5.6	1277	4	US-09-397-885-3
6	89.5	5.5	387	1	US-08-112-630-2
7	88.5	5.5	1529	4	US-09-134-001C-3945
8	88	5.4	1161	4	US-09-327-536-2
9	86.5	5.4	428	4	US-09-548-372D-51
10	86.5	5.4	428	4	US-09-548-367D-51
11	86.5	5.4	434	4	US-09-548-372D-53
12	86.5	5.4	434	4	US-09-548-367D-53
13	86.5	5.4	476	4	US-09-548-372D-6
14	86.5	5.4	476	4	US-09-548-372D-73
15	86.5	5.4	476	4	US-09-548-367D-6
16	86.5	5.4	476	4	US-09-548-367D-73
17	86.5	5.4	526	4	US-09-342-648-8
18	84	5.2	1112	2	US-08-714-402-2
19	83.5	5.2	409	4	US-09-134-001C-3790
20	83	5.1	758	4	US-09-134-001C-4588
21	81.5	5.0	568	1	US-08-528-199-1
22	81.5	5.0	1579	3	US-08-755-587-184
23	81	5.0	287	3	US-09-173-581-2
24	81	5.0	287	4	US-09-420-915-2
25	81	5.0	288	4	US-09-441-039-2
26	80.5	5.0	409	4	US-09-071-035-316
27	80.5	5.0	433	4	US-09-071-035-314

Sequence 16, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 58, Appl
Sequence 58, Appl
Sequence 2, Appl
Sequence 45, Appl
Sequence 4705, Ap
Sequence 1, Appl
Sequence 9, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 1, Appl
Sequence 4, Appl
Sequence 4, Appl

US-08-755-587-16
US-08-747-221B-14
US-09-005-051-14
US-08-747-221B-58
US-09-005-051-58
US-09-342-648-2
US-08-887-534A-45
US-09-134-001C-4705
US-08-965-600-1
US-09-489-506-1
US-08-836-854-9
US-08-747-221B-53
US-09-005-051-53
US-08-747-221B-19
US-09-005-051-19
US-09-393-627B-1
US-07-961-522-4
US-08-217-438-4

ALIGNMENTS

RESULT 1

US-09-134-001C-4646

; Sequence 4646, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4646

; LENGTH: 332

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-4646

Query Match 6.8% Score 110; DB 4; Length 332;
Best Local Similarity 24.5%; Pred No. 0.00084;
Matches 58; Conservative 40; Mismatches 111; Indels 28; Gaps 11;

QY 7 KIAELGKYTVGEGPHWDHETQTLVFDVTEKTHKYPVPSQKKYTFCKVDKLVSFIIPLAG 66

Db 44 KISKDGLQL--EGLNFNREGQ-LFLDDVFEGNFKVNPATKEVT---TKFQS---VKD 92

QY 67 SPGRFVVSLEREIAIL-----TWGVSAAPTSTEAIVNVEPHIKNNR-LNDGKADPLGN 119

Db 93 NPAIKVHKDGRFLCYLGDFKTFATGEQIEEIIISDLNTEYCDMDVDFSKGG 152

QY 120 LWTGTMAIDAGLPVPTGSLYHLGAD-KVKMHESNIAIANGLAWSNDLKKMYIDSGK 178

Db 153 FYFTDF---RGYSTOPL-GGVYYVDPDFKVTPTTIQNISVANGIALSTDEKVLVWTTT 208

QY 179 RVVDE--YDYDASTLSISNORPLTFTEKHEVPGPDGQTDIEGNLWAVFQGOIII 233

Db 209 NRLHRLALEDGVTIAPFGATIPYFTTGHG---GPDSCCDSNDNLVYAMYGQGRVL 262

RESULT 2

US-09-134-001C-5111

; Sequence 5111, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GFC-007
CURRENT APPLICATION NUMBER: US/09/134, 001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5111
LENGTH: 413
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5111

Query Match 5.8%; Score 94; DB 4; Length 413;
Best Local Similarity 22.6%; Pred. No. 0.074;
Matches 76; Conservative 46; Mismatches 101; Indels 114; Gaps 17;

QY 36 EKTTHKY-VPSQKKYTCCKVDKLVSLFIPLAGSPGRFV-----SLEREIAL-- 82
DB 6 EKNQHYRNSQKRHTFPWIKTIIVAI--AGIIGALLVIGICKLLNKTGFNNEGATVHQ 63
QY 83 -----TWGVSAAFTSI-EATVNVPEHIKNRNLNDGKADPLGNLWTG--TMAIDAGL 131
DB 64 VNSHGGNQLDGKSNQYKSVHDMIKOVSPAIV-GVINMQKSTNLDLDFNGKASKSKEAGI 122
QY 132 PVGPV-----TGSLY-----HL--GADK-KVKMHESN----- 155
DB 123 GSGVIYQISDGSAYIVTNNHVDGASEIKVLQHSKOVDAKLGKDALTDIAVLKTKDKT 182
QY 156 -----IATAN-----GLAWSNLDKMYIDSGKRRVDEYDASTLSIS 194
DB 183 GIKAIQFANSKVKVOTGDSVFAMGNPLGLEFANS-----VTSGIISASERTIDANTSAGN 236
QY 195 NQRLPFTFEKHEVPGYDPGOTIDEEGNLWAVFOGRIIKISTQOPEVLDTVKIPDPQV 254
DB 237 TKVNVLOTDAAINFGNSGGALVDINGNL-----VGINSMKIRAAQV 277
QY 255 TSAVFGPNLDLHVTSLAGL-----OLDDSSLDKSLVN 287
DB 278 EGIGFAIPS-NEVRVTIEQLVKHGKIERSIGIGLIN 313

RESULT 3

US-08-501-126-2
Sequence 2, Application US/08501126
Patent No. 6140095
GENERAL INFORMATION:
APPLICANT: Van Solingen, Pieter
APPLICANT: Williams, Diane P.
APPLICANT: Iverson, Sara
APPLICANT: Farrell, Roberta L.
APPLICANT: Herbes, Wilhelmina T.
APPLICANT: Van Der Kleij, Wilhelmus A.
APPLICANT: Herweijer, Margaretha A.
APPLICANT: Van Beckhoven W.C., Rudolf F.
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Jones, Brian E.
TITLE OF INVENTION: ALKALI-TOLERANT XYLANASES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/501,126
FILING DATE: 29-DEC-1995
CLASSIFICATION: 425
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615-0057.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-501-126-2

Query Match 5.8%; Score 93; DB 4; Length 396;
Best Local Similarity 24.7%; Pred. No. 0.089;
Matches 60; Conservative 36; Mismatches 83; Indels 64; Gaps 15;
QY 100 VEPHIKNRNLNDGKADPLGNLWTGTMADAGLPVG--PVTGSLYHLGADKKVKM-HESNI 156
DB 71 VEPY-----OLEGRQAOLKHHYNSLVAENAMKPVSLQPREGEMWEGADKIVEFARKHNM 126
QY 157 ATA-NGLSNLDLKKMYIID-SGKRRVDEYDASTLSISNORPLTFTEKH---EVPGY 211
DB 127 ELRFHTLVHMSQVPEFFIDENGNRVDETDPEKR--KANKQLLLERMMENHIKTVVVERVK 184
QY 212 D-----GQIDEEGNL-----WVAVFOGRIIKI-----STOO 239
DB 185 DDVTSMDVNVNEVIDDGGGLRESEWYQI-TGTDYIKVAFETARKYGGEEAKLYINDYNTVE 243
QY 240 P---EVLLDTVK-----IPDPQV---TSVAGGPNLDLHV-----TSAGLQDDSSLD 282
DB 244 PSKRDDLYNLVKDLLEGGVPIDCGVGHQSHIOGWPSIEDTRASFETSLGLDNQVTELD 303
QY 283 KSL 285
DB 304 MSL 306

RESULT 4

US-08-282-197C-55
Sequence 55, Application US/08282197C
Patent No. 5871730
GENERAL INFORMATION:
APPLICANT: Brzezinski, Ryszard
APPLICANT: Dery, Claude V
APPLICANT: Beaulieu, Carole
TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and
METHODS OF USE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave., NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,197C
FILING DATE: 29-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A

REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1050.0410000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-282-197C-55

Query Match 5.6%; Score 91; DB 2; Length 397;
Best Local Similarity 24.7%; Pred. No. 0.15;
Matches 60; Conservative 35; Mismatches 84; Indels 64; Gaps 15;

100 VEPHIKNNRLNDGKADPLGNLTGTMAIDAGLP--VGPVTGSLYHLGADKKVKM--HESNI 156
171 VEPY-----OLEGROAQILKHYNLSVAENAKPESLQPREGEWNWEGADKIVEFARKHNM 126
157 AIA-NGLAWSNDLKKMYID--SGRRVDEYDYDASTLSISNORPLFTFEKH---EVPGY 211
127 ELRPHTLVHWSQVPEWFEIDEDGNRMVDETDPKR--EANKQLLLERMEHIKTIVVERK 184
212 D-----GQIDEEGNL-----WVAVFQGGRIKI-----STQQ 239
185 DDVTSWVNVNEVIDDGGGLRESEWYQI--TGTDYIKVAFETARKYGGEEAKLYINDYNTEV 243
240 P---EVLIDTVK-----IPDPQV---TSVAFGGPNLDLHV-----TSAGLQDDSLD 282
244 PSKRDLYNLVKDLLEQGVPIDGCHOSHIOIGWPSIEDTRASFEKFTSLGLDNQVTELD 303
283 KSL 285
304 MSL 306

RESULT 5
US-09-397-885-3
Sequence 3, Application US/09397885
Patent No. 6323007
GENERAL INFORMATION:
APPLICANT: Moller, Soren
APPLICANT: Johansen, Charlotte
APPLICANT: Schafer, Thomas
APPLICANT: Ostergaard, Peter Rahbek
APPLICANT: Hoeck, Lisbeth Hedegaard
TITLE OF INVENTION: A 2,6,-b-D-Fructan Hydrolase Enzyme And
TITLE OF INVENTION: Process For Using The Enzyme
FILE REFERENCE: 5540.200-US
CURRENT APPLICATION NUMBER: US/09/397.885
CURRENT FILING DATE: 1999-09-17
EARLIER APPLICATION NUMBER: PA 1998 01173
EARLIER FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: PA 1998 01623
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 60/101.615
EARLIER FILING DATE: 1998-09-24
EARLIER APPLICATION NUMBER: 60/111.675
EARLIER FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1277
TYPE: PRT
ORGANISM: Paenibacillus pabuli
US-09-397-885-3

Query Match 5.6%; Score 90; DB 4; Length 1277;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 55; Conservative 44; Mismatches 102; Indels 58; Gaps 12;

664 MPNORKVNLAWMTNDYFPEPTSSWKGQ--LTIPREVSLRTTDEGVRVLTQPTITELQKL 721
98 -----VNVPEHIKNNRLNDGKADPLGNLTGTMAIDAGLPVGPVTGSLYHLGADKK 148
722 RHNLYSAQOQWTVGPKSKN-----PLEGLTAGAYEIEAEVEI--PANSSVTEGFOUR 771
149 VKMHESNTAIANGLAWSNDLKKMYIDSGKRRVDEYDYDASTLSISNORPLFTFEKHEVP 208
772 QREGOKTT-----VAYRVDTQNM-----VDRTTSGDVFSDLFT--KVHEAS 812
209 GYPGQOTIDEEGNLWVAVFQGGRIKIKISTQPEVLLDTVKIPDPQVTSVAF---GGPNLD 265
813 LKPNQKVK-----LRIFVDESSVEFEGNDGKVVFSDVIPDPAGRAMAFYSLGG----- 862
266 ELHVTISAGLQDDSLDSSDKS 284
863 EVKYSKMYALDNIWRKS 881

RESULT 6
US-08-112-630-2
Sequence 2, Application US/08112630
Patent No. 5411886
GENERAL INFORMATION:
APPLICANT: UDAKA, Shihezo
APPLICANT: SAKAGUCHI, Kenji
APPLICANT: YAMAGATA, Hideo
APPLICANT: DEKKER, Koen
TITLE OF INVENTION: XYLOSEISOMERASE GENE OF THERMUS
TITLE OF INVENTION: AQUATICUS, XYLOSEISOMERASE AND PROCESS FOR PREPARATION OF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker and Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/112.630
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/783,150
FILING DATE: 28-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Rea, Teresa S.
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 024705-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
TELEX: 440 580
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-112-630-2

Query Match 5.5%; Score 89.5; DB 1; Length 387;
Best Local Similarity 19.8%; Pred. No. 0.21;
Matches 74; Conservative 57; Mismatches 146; Indels 97; Gaps 15;

4 VVERIAELGKTYVG-----EGPHWDHETQTLTYFVDTVEKTHKYV-----PS 45

Db 38 VVHKLAEALGAYGVNLHDEDLIPRGTPPOERDOIVRRFKKALDETLGLKVPWVTANLFSDDPA 97
QY 46 QKKYTFCKVVKLV-----FIPLAGSPGRFVVSLE-----REI 79
Db 98 FKGAFSPDPWVRAYALRKSLEIMDLGAEIGAELIYVV-----WPCREGAEVATGKARKV 153
QY 80 AILTWGVSAAPTSTEAIVNVEPHIKNNRLNDGKADPLGNLWGTMTADAGLPVGPVTS 139
Db 154 ----WDWREALNEMAAEAEDQGYRFALEPKPNEPRGDIYFAT-----VGSMLAF 201
QY 140 LYHLGADKKVK-----HE--SNIAIANGLAWSNDLKKMYIDSGKRRVDEYDASTLS 192
Db 202 IHTLDRPERGLNPEFAHETWAGLNFVHVAQAALDAGKLFDHLDNDQMSRFDQDLRGS 261
QY 193 ISNORPLTFEKEHVPYGPDG-----QTIDEENLWAVFQOQRIIKISTOQPEVLL 244
Db 262 ENLKAFFLVLDLESSGYGPRHEDAHALRTEDEBG-VWAFARGCMRTYLLIKERAFAFR 320
QY 245 DTVKIPDPQV-----TSVAFGPNLDELHVTISAGLQDLDSSLDKSLYNGH-VY 291
Db 321 E-----DPEVKELLAAYQEDPAALALLGPSRE---KAEALKRAELPLEAKRRRGYALE 372
QY 292 RVTGLGVKGAGVK 305
Db 373 RLQLAVEYLLGVR 386

RESULT 7
US-09-134-001C-3945
; Sequence 3945, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3945
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3945

Query Match 5.5%; Score 88.5; DB 4; Length 1529;
Best Local Similarity 22.8%; Pred. No. 2.6;
Matches 54; Conservative 29; Mismatches 87; Indels 67; Gaps 9;

QY 50 TFCKVDKLVFIPLAGSPGRFVVSLEIREIALITWGVSAAPTSTEAIVNVEPHIKNNRL 109
Db 375 SFCNGKIGALTRNGLRGRVITTKDNFVFSSEVGVIDVPEENVAF-----KGOL 426
QY 110 NDGK---ADPLGNLWGTMAIDAGLPVPGVTSGLYHLGADKKVKHESNIAIANGLA-- 164
Db 427 NPGKLLLVDFLQN-----KVVENNELKTNLANELPYEQ 459
QY 165 -----SNDLKKMYIDSGKRRVDEYDASTLSISNORPLTFEKEHVPYGPDGQTID 217
Db 460 WLKDYKNKNDLNIYQSS-----DWDDQTLFRLQK--FAYTKEDINKYMTDLVIN 509
QY 218 EGNLWAVFQOQRIIKISTOQPEVLLDTVTKIPDPQVTSVARGPNLD---ELHVTIS 271
Db 510 KKDPIGAMGYDAP--IAVLNDKPSLEFNFKQLFAQVTN-----PPIDAYREKIVTS 559

RESULT 8
US-09-327-536-2

; Sequence 2, Application US/09327536
; Patent No. 6355477
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; APPLICANT: ROCHA, Claudia
; TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
; TITLE OF INVENTION: STREPTOCOCCI
; FILE REFERENCE: 022927-008
; CURRENT APPLICATION NUMBER: US/09/327,536
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: US 08/714,402
; PRIOR FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: SFFBP gene
; US-09-327-536-2

Query Match 5.4%; Score 88; DB 4; Length 1161;
Best Local Similarity 17.5%; Pred. No. 1.9;
Matches 64; Conservative 55; Mismatches 122; Indels 124; Gaps 16;

QY 13 KYTVGEGPHWDHETQTLFYVDTEKTFHKYVPSQKRYTFCKVDKLVFIIPLAGSPGRFV 72
Db 780 KFISSGNEWSFEFKNL-----KKY-----NGTGNDI 806
QY 73 VSLEREIAI-----LTWGVSAAPTSTEAIVNVEPHIKNNR---LNDGKADPLGNLWGT 124
Db 807 IYSVKEVTPVTGYDVTYSANDIINTKREVIITQOGPKLEIETPLPESGASG-----GT 859
QY 125 MAIDAGLPVPGVTSGLYHLGADKKVKHESNIAIANGLAWSNDLKKMYIDSGKRRVDEY 184
Db 860 TTVEDSRPDTLSGLSSQSGSDMTTEEDSAT-----HIKFSKRDIDGK 904
QY 185 DYDASTLSI--SNORPLTF-----EKHEVPG-----YDQG-----TDEE 219
Db 905 ELAGATMELRDSGKTIISTWISDGVKDFYLMPPGKYFETVETAAPDGYBIATAITFTVNEQ 964
QY 220 GNLWV---AVFQOQRIIKISTQCP-----EVLDD-----TVKIPDPQVTSVA 258
Db 965 GOVTYNGKATGDTTHIVNMVDAYKPKGSGQVIDIEEKLDPDEQHGSGSTTEIEDSKSLLI 1024
QY 259 FGGPN--LDELHVTISAGL-----QLDSSLSOKSLVNGHYRV-----TCLGVKGPAG 303
Db 1025 IGGQGEVVDTTEDTQSGMTGHSSTTEIEDSKSDVIICGGQGVVETTEDTQTGMYGDSG 1084
QY 304 VVKVL 308
Db 1085 CKTEV 1089

RESULT 9
US-09-548-372D-51
; Sequence 51, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Hu-Asp2(b)
US-09-548-372D-51

```

Query Match	5.4%	Score	86.5	DB	4	Length	428
Best Local Similarity	22.0%	Pred.	No.	0.54			
Matches	63	Conservative	31	Mismatches	95	Indels	97
						Gaps	15
QY	2	GPVVEKTAEL	-----GKY--	IVGEGPHDHD	HTQTL	YFVDVT	VEKTF-----HK
			:		:		:
Db	58	GSFVEMVDN	LRGSGGQGYVEM	VGSP	-----QTLN	ILVDTGSSNF	AVGAAPHLHR
			:		:		:
QY	42	YVFSQKYTECK	VDKLVFSFIPLA	-----GSPGR	FVVSLE	REIAILL	TWDCVSAAPTSEA
			:		:		:
Db	112	YYORQLSST	TRDLKKGV	--VPY	TQGWEGEL	GPLDVS	IPHGNPTVRANIA
			:		:		:
QY	97	IVNVEPHIKN	NRLNDGRADPL	GNLTWT	CTMAI	-----DAGLP	-----VGPVTGS
			:		:		:
Db	170	FIN	-----GSN	WEGILGLAY	AEIARLC	GAGPLNQ	SEVLASVGS
			:		:		:
QY	142	HLGADKVK	KMHESNIA	IANG	LAWSND	LKMYIID	SGRRVD-----EYD
			:		:		:
Db	212	IGGDH	-----SLYT	GSMLWTP	PIRREWY	EVIIVR	VEINGQDLK
			:		:		:
QY	191	L--SIS	NQR-PLTTFE	-----KHE	VPYDP	GQTI	DEBGNLWA
			:		:		:
Db	263	VDSGTTN	LRPKPKVFE	AAVKS	SIKAAS	STKFP	DPGDFWLG
			:		:		:

```

Query Match      5.4%; Score 86.5; DB 4; Length 428;
Best Local Similarity 22.0%; pred. No. 0.54;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVEKTAEL-----CKY---TVGEGPHNDHETQL-YFVDVTEKTF-----HK 41
Db 58 GSFVEVDNLRGSGQGYVEMTVGSP-----QTLNILDVGSSNFVGAAPHFVLR 111
QY 42 YVPSQKKYTECKVDKLSFYIPLA-----GSPGRFVVSLEIREAILTWGDVSAAPTSEA 96

```

```
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 3.1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Hu-Asp2(b) delta TM
US-09-548-367D-53

Query Match          5.4%; Score 86.5; DB 4; Length 434;
Best Local Similarity 22.0%; Pred. No. 0.55;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVEKIAEL-----GKY---TVGEGPHMDHETQTL-YFVDIVTEKTF-----HK 41
DB 58 GSFVEMVDNLKSGKSGQGYVEMTVGSP-----QTLNILDVDTGSSNFVAGVAAAPHPFLHR 111
QY 42 YVPSQKKYTFCKVKDKLVSFIIPLA-----GSPGRFVVSLEIREIAILTWDGVSAAAPTISIA 96
DB 112 YYQRLSSITYRDLKRGV--YVPYTGKWEGLGTLVSIPIHGPNTVTRANIAAITESDKF 169
QY 97 IVNVEPHIKNNRLNDGKADPLGNLTGTMAI-----DAGLP-----VGPVTGSLY 141
DB 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211
QY 142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSCKRRVD-----EYDYDAST 190
DB 212 IGGIDH-----SLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSI 262
QY 191 L--SISNOR-PLFTFE-----KHEVPGYPDGTIDEENLAWA 225
DB 263 VDSGTTNLRPKKVFEEAAVKSIAASSTEKFPDGFGLGELQVLCWQA 308

RESULT 13
US-09-548-372D-6
; Sequence 6, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-73

Query Match          5.4%; Score 86.5; DB 4; Length 476;
Best Local Similarity 22.0%; Pred. No. 0.64;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVEKIAEL-----GKY---TVGEGPHMDHETQTL-YFVDIVTEKTF-----HK 41
DB 58 GSFVEMVDNLKSGKSGQGYVEMTVGSP-----QTLNILDVDTGSSNFVAGVAAAPHPFLHR 111
QY 42 YVPSQKKYTFCKVKDKLVSFIIPLA-----GSPGRFVVSLEIREIAILTWDGVSAAAPTISIA 96
DB 112 YYQRLSSITYRDLKRGV--YVPYTGKWEGLGTLVSIPIHGPNTVTRANIAAITESDKF 169
QY 97 IVNVEPHIKNNRLNDGKADPLGNLTGTMAI-----DAGLP-----VGPVTGSLY 141
DB 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211
QY 142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSCKRRVD-----EYDYDAST 190
DB 212 IGGIDH-----SLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKSI 262
QY 191 L--SISNOR-PLFTFE-----KHEVPGYPDGTIDEENLAWA 225
DB 263 VDSGTTNLRPKKVFEEAAVKSIAASSTEKFPDGFGLGELQVLCWQA 308

RESULT 14
US-09-548-372D-73
; Sequence 73, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-73

Query Match          5.4%; Score 86.5; DB 4; Length 476;
Best Local Similarity 22.0%; Pred. No. 0.64;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVEKIAEL-----GKY---TVGEGPHMDHETQTL-YFVDIVTEKTF-----HK 41
DB 58 GSFVEMVDNLKSGKSGQGYVEMTVGSP-----QTLNILDVDTGSSNFVAGVAAAPHPFLHR 111
QY 42 YVPSQKKYTFCKVKDKLVSFIIPLA-----GSPGRFVVSLEIREIAILTWDGVSAAAPTISIA 96
DB 112 YYQRLSSITYRDLKRGV--YVPYTGKWEGLGTLVSIPIHGPNTVTRANIAAITESDKF 169
QY 97 IVNVEPHIKNNRLNDGKADPLGNLTGTMAI-----DAGLP-----VGPVTGSLY 141
DB 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211
QY 142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYIDSCKRRVD-----EYDYDAST 190
```

```

Db      212 IGGIDH-----SLYTGSLWTPIRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
Qy      191 L--SISNQR-PLTFE-----KHEVPGYPDGQTTIDEENLWVA 225
Db      263 VDSGTTNRLPKKVFEEAAVKSIRKASSTKFPDGFGLGEQLVCWQA 308

RESULT 15
US-09-548-367D-6
; Sequence 6, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-6

Query Match      5.48; Score 86.5; DB 4; Length 476;
Best Local Similarity 22.08; Pred. No. 0.64;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

Qy      2  GPVVEKIAEL-----GKY---TVGEGPHWDHETOTL-YFVDTVEKTF-----HK 41
Db      58  GSFVEMVDNLRGKSGQYVEMTVGSP-----QTLNILDVTGSSNFVGAAPHPLHR 111
Qy      42  YVPSQKKYTCVKDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWGVSAAPTSIEA 96
Db      112 YVORQLSSTYRDLRKGV--VVPYTGKWEGELGTDLSIPHPNVTVRANIAAITESDKF 169
Qy      97  IVNVEPHIKNRLNDGRADPLGNLWTGTMAI-----DAGLP-----VGPVTGSLY 141
Db      170 FIN-----GSNWEGILGLAYAEIARLCGAGPPLNOSEVLASVGGSMI 211
Qy      142 HLGADKKVKMHESNIAIANGIANSNDLKKMYIDSGRKRVY-----EYDYDAST 190
Db      212 IGGIDH-----SLYTGSLWTPIRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
Qy      191 L--SISNQR-PLTFE-----KHEVPGYPDGQTTIDEENLWVA 225
Db      263 VDSGTTNRLPKKVFEEAAVKSIRKASSTKFPDGFGLGEQLVCWQA 308
```

Search completed: August 1, 2003, 14:35:38
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:35:09 ; Search time 47 Seconds
(without alignments)
778.256 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGVPVKIAELGKTYVGGP.....HVVRYTGLGVKGAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCST_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCSTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep2.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep3.*
- 13: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	311.5	19.3	286	15	US-10-156-761-13272
2	92	5.7	4349	15	US-10-160-758-15
3	90	5.6	1277	10	US-09-969-362-3
4	89.5	5.5	418	15	US-10-156-761-12646
5	88.5	5.5	610	11	US-09-738-626-4495
6	87	5.4	450	15	US-10-156-761-14443
7	86.5	5.4	467	11	US-09-738-626-6671
8	86.5	5.4	428	10	US-09-794-927-51
9	86.5	5.4	428	10	US-09-795-847-51
10	86.5	5.4	428	10	US-09-794-743-51
11	86.5	5.4	428	10	US-09-794-748-51
12	86.5	5.4	428	10	US-09-794-925-51
13	86.5	5.4	428	10	US-09-681-442-51
14	86.5	5.4	428	12	US-09-869-414-51
15	86.5	5.4	428	12	US-09-548-366-51
16	86.5	5.4	434	10	US-09-794-927-53

17	86.5	5.4	434	10	US-09-795-847-53	Sequence 53, Appl
18	86.5	5.4	434	10	US-09-794-743-53	Sequence 53, Appl
19	86.5	5.4	434	10	US-09-794-748-53	Sequence 53, Appl
20	86.5	5.4	434	10	US-09-794-925-53	Sequence 53, Appl
21	86.5	5.4	434	10	US-09-681-442-53	Sequence 53, Appl
22	86.5	5.4	434	12	US-09-869-414-53	Sequence 53, Appl
23	86.5	5.4	434	12	US-09-548-366-53	Sequence 53, Appl
24	86.5	5.4	476	10	US-09-794-927-6	Sequence 6, Appl
25	86.5	5.4	476	10	US-09-795-847-6	Sequence 6, Appl
26	86.5	5.4	476	10	US-09-794-743-6	Sequence 6, Appl
27	86.5	5.4	476	10	US-09-794-748-6	Sequence 6, Appl
28	86.5	5.4	476	10	US-09-794-925-6	Sequence 6, Appl
29	86.5	5.4	476	10	US-09-681-442-6	Sequence 6, Appl
30	86.5	5.4	476	12	US-09-869-414-6	Sequence 6, Appl
31	86.5	5.4	476	12	US-09-548-366-6	Sequence 6, Appl
32	86.5	5.3	393	14	US-10-080-170-166	Sequence 15, Appl
33	84.5	5.2	1139	12	US-09-820-843A-15	Sequence 15, Appl
34	83.5	5.2	770	15	US-10-281-478-11	Sequence 11, Appl
35	83	5.1	340	11	US-09-738-626-6960	Sequence 6960, Ap
36	82.5	5.1	462	15	US-10-156-761-11331	Sequence 11331, A
37	82.5	5.1	527	11	US-09-934-868-60	Sequence 60, Appl
38	82.5	5.1	4545	10	US-09-873-403-2	Sequence 2, Appl
39	82	5.1	3472	15	US-10-027-806-4	Sequence 4, Appl
40	82	5.1	3472	15	US-10-034-623-4	Sequence 4, Appl
41	82	5.1	3472	15	US-10-027-801-4	Sequence 4, Appl
42	81.5	5.0	1294	10	US-09-815-242-13724	Sequence 13724, A
43	81.5	5.0	2042	15	US-10-192-584-6	Sequence 6, Appl
44	81	5.0	287	10	US-09-870-962-2	Sequence 2, Appl
45	81	5.0	288	15	US-10-165-696-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-10-156-761-13272
; Sequence 13272, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13272
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13272

Query Match 19.3%; Score 311.5; DB 15; Length 286;
Best Local Similarity 32.6%; Pred. No. 1.6e-23;
Matches 84; Conservative 45; Mismatches 112; Indels 17; Gaps 8;
Qy 16 VGEQPHWDHETQTLFYVDVTEKTHKYVPSOKKTFCKDKLVFIILPAGSPGRFVSL 75
Db 14 LGEQPTWTAORLIWDILGSRVHTYDPSGRRTVLATEQHVGAAPRAG--GGLVNNL 71
Qy 76 EREAITLWDGVSAAPTSTIEALVNVPEPHKNNRLNDGKADPLGNLWTGTMAIDAGLPVGP 135
Db 72 RDGVLGYPEGPGGRDF---RWLHREP--VPGRRGNDAAVAADGSLWAGTMYRDE---A 123

; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 12646
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-12646

Query Match 5.5%; Score 89.5; DB 15; Length 418;
 Best Local Similarity 26.5%; Pred. No. 1.5;
 Matches 26; Conservative 17; Mismatches 40; Indels 15; Gaps 3;

QY 44 PSOKKYTFCKVCKLVSFIIPLAGSPGRFVSVLEREIAILTWDCVSAAPTSTEAIVNVEPH 103
 DB 246 PARRYPWCLTMAALLIPLAA--RLTGLLALAGALLAGMATAPMTWMTFLVOOR 303
 QY 104 IKNNRLDKADPLGNLTGWTMAID--AGLPVGPVGT 138
 DB 304 TPESRLNEGM-----TLAVTGLLGIACGSATG 331

RESULT 5
 US-09-738-626-4495
 ; Sequence 4495, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4495
 ; LENGTH: 610
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4495

Query Match 5.5%; Score 88.5; DB 11; Length 610;
 Best Local Similarity 21.7%; Pred. No. 3.4;
 Matches 70; Conservative 48; Mismatches 89; Indels 115; Gaps 20;

QY 14 YTVGEGPHW----DHETOTLYFVDVVEKTEHYV-PSOKKYTFCKVCKLVSFIIPLAGS 67
 DB 272 YRSGDPEARWTFNDPETHSYFNGKDNITFHSQIWPAB-----LLGYAGK 317
 QY 68 PGR-----FVSVLEREIA--ILTWDCVSAAPTSTEAIVNVEPHIK----- 105
 DB 318 GSRGGEIGDLGVLNLTPEVVSSEFLTMSG-SKFSKSGVVIIVKDFLKEFGPDALRYFIA 376
 QY 106 -----NN-----RLNDGKADPLGNLTGWTMAIDAGLPVG--PVTGSLYHLGADK 147
 DB 377 VAGPENNDTFTWDFEVRVNNELANGWGNLVRTVSM-AHKNFGEVVPVPGALEE--SDK 433
 QY 148 KVK-----MH-----ESNTAINGLAWSNLDLKKMYIDSGK 178
 DB 434 KILDATAFESVAANLDOSKFKAGISEIMHVVGANAYIAQEPW-----KLAKDDTKR 488

; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 12646
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-12646

Query Match 5.5%; Score 89.5; DB 15; Length 418;
 Best Local Similarity 26.5%; Pred. No. 1.5;
 Matches 26; Conservative 17; Mismatches 40; Indels 15; Gaps 3;

QY 44 PSOKKYTFCKVCKLVSFIIPLAGSPGRFVSVLEREIAILTWDCVSAAPTSTEAIVNVEPH 103
 DB 246 PARRYPWCLTMAALLIPLAA--RLTGLLALAGALLAGMATAPMTWMTFLVOOR 303
 QY 104 IKNNRLDKADPLGNLTGWTMAID--AGLPVGPVGT 138
 DB 304 TPESRLNEGM-----TLAVTGLLGIACGSATG 331

RESULT 5
 US-09-738-626-4495
 ; Sequence 4495, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4495
 ; LENGTH: 610
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4495

Query Match 5.4%; Score 87; DB 15; Length 450;
 Best Local Similarity 22.1%; Pred. No. 3;
 Matches 38; Conservative 23; Mismatches 51; Indels 60; Gaps 7;

QY 111 DGRADPLGNLTGWTMAID---AGL-----PVGPTGSLYHLGADKVKMHE 153
 DB 163 DGQSDGKSSIWDTFAIDDTASGLRLVSYRLTLRTPGGTTPTVWRIGA----- 214
 QY 154 SNIAIANGLAWSNLDLKKMYIDSGKRRVDEYDASTLSISNQR--PLFTFEKH--EVPG 209
 DB 215 -----MGSDIPDRFTVPASAPGLAQELRVPKYSQEIHAQOYPE 252
 QY 210 YPDGQTIDEENGLMWAVFGQRIKISTQOP-----EVLDDTVKIPDPQVTSVA 258
 DB 253 YDNG-----GEAWCSPTSSQMIIEYWRRTPTQGLAWVDPVSADPQVCHAA 298

RESULT 7
 US-09-738-626-6671
 ; Sequence 6671, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484

; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 12646
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-12646

Query Match 5.5%; Score 89.5; DB 15; Length 418;
 Best Local Similarity 26.5%; Pred. No. 1.5;
 Matches 26; Conservative 17; Mismatches 40; Indels 15; Gaps 3;

QY 44 PSOKKYTFCKVCKLVSFIIPLAGSPGRFVSVLEREIAILTWDCVSAAPTSTEAIVNVEPH 103
 DB 246 PAERYPNCWLTMAALLIPLAA--RLTGSLLAGALLAGMATAPTMTAMTLVVOOR 303
 QY 104 IKNNRDKGADPLGNLTGWTMAID--AGLPVGPVGT 138
 DB 304 TPESRLNEGM-----TLAVTGLLGIACGSATG 331

RESULT 5
 US-09-738-626-4495
 ; Sequence 4495, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4495
 ; LENGTH: 610
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4495

Query Match 5.5%; Score 88.5; DB 11; Length 610;
 Best Local Similarity 21.7%; Pred. No. 3.4;
 Matches 70; Conservative 48; Mismatches 89; Indels 115; Gaps 20;

QY 14 YTVGEGPHW----DHETOTLYFVDVVEKTEHYV-PSOKKYTFCKVCKLVSFIIPLAGS 67
 DB 272 YRSGDPEARWTFNDPETHSYFNGKDNITFHSQIWPAB-----LLGYAGK 317
 QY 68 PCR-----FVSVLEREIA--ILTWDCVSAAPTSTEAIVNVEPHIK----- 105
 DB 318 GSRGGEIGDLGVLNLTPEVVSSEFLTMSG-SKFSKSGVVIIVKDFLKEFGDRLRYFIA 376
 QY 106 -----NN-----RLNDGKADPLGNLTGWTMAIDAGLPVG--PVTGSLYHLGADK 147
 DB 377 VAGPENNDTFTWDFVFRVNNELANGWGNLVRTVSM-AHKNFGEVVPVGALEE--SDK 433
 QY 148 KVK-----MH-----ESNTAINGLAWSNLDLKKMYIDSGK 178
 DB 434 KILDLTAFESVAANLDOSKFKAGISEIMHVVGANAYIAQEPW-----KLAKDDTKR 488

; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 12646
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-12646

Query Match 5.5%; Score 89.5; DB 15; Length 418;
 Best Local Similarity 26.5%; Pred. No. 1.5;
 Matches 26; Conservative 17; Mismatches 40; Indels 15; Gaps 3;

QY 44 PSOKKYTFCKVCKLVSFIIPLAGSPGRFVSVLEREIAILTWDCVSAAPTSTEAIVNVEPH 103
 DB 246 PAERYPNCWLTMAALLIPLAA--RLTGSLLAGALLAGMATAPTMTAMTLVVOOR 303
 QY 104 IKNNRDKGADPLGNLTGWTMAID--AGLPVGPVGT 138
 DB 304 TPESRLNEGM-----TLAVTGLLGIACGSATG 331

RESULT 5
 US-09-738-626-4495
 ; Sequence 4495, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4495
 ; LENGTH: 610
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4495

Query Match 5.5%; Score 88.5; DB 11; Length 610;
 Best Local Similarity 21.7%; Pred. No. 3.4;
 Matches 70; Conservative 48; Mismatches 89; Indels 115; Gaps 20;

QY 14 YTVGEGPHW----DHETOTLYFVDVVEKTEHYV-PSOKKYTFCKVCKLVSFIIPLAGS 67
 DB 272 YRSGDPEARWTFNDPETHSYFNGKDNITFHSQIWPAB-----LLGYAGK 317
 QY 68 PCR-----FVSVLEREIA--ILTWDCVSAAPTSTEAIVNVEPHIK----- 105
 DB 318 GSRGGEIGDLGVLNLTPEVVSSEFLTMSG-SKFSKSGVVIIVKDFLKEFGDRLRYFIA 376
 QY 106 -----NN-----RLNDGKADPLGNLTGWTMAIDAGLPVG--PVTGSLYHLGADK 147
 DB 377 VAGPENNDTFTWDFVFRVNNELANGWGNLVRTVSM-AHKNFGEVVPVGALEE--SDK 433
 QY 148 KVK-----MH-----ESNTAINGLAWSNLDLKKMYIDSGK 178
 DB 434 KILDLTAFESVAANLDOSKFKAGISEIMHVVGANAYIAQEPW-----KLAKDDTKR 488

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6671
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6671

Query Match 5.4%; Score 87; DB 11; Length 467;

Best Local Similarity 35.8%; Pred. No. 3.2;
Matches 29; Conservative 10; Mismatches 26; Indels 16; Gaps 5;

Qy 81 ILTWDGVSAAPTSEIAIVNVEPHI--KNNRLNDGKADPLGNLTGTWTAIDAGLPVGPVTG 138
Db 391 IRSW--LDGGPTDENTNICPHDHGNNNDQDQK-DNWGH-----MNID-----PTTG 436
Qy 139 SLYHLGADKKVKMHESNTAIA 159
Db 437 RVGQPADRRKPMRFNNTAAA 457

RESULT 8

US-09-794-927-51
; Sequence 51, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-794-927-51

Query Match 5.4%; Score 86.5; DB 10; Length 428;

Best Local Similarity 22.0%; Pred. No. 3.2;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

Qy 2 GPVVEKIAEL-----GKY---TVGEGPHDHTQTL-YFVDTVKTF-----HK 41
Db 58 GSFVEMVDNLKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAVGAAPHELHR 111
Qy 42 YVPSQKYYTFCKVKLVSIPIA-----GSPGRFVSLEREIALITWDGVSAAPTSEA 96

Db 112 YQORLSSTYRDLRKG--YVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAAITESDKF 169
Qy 97 IVNVEPHIKNNRLNDGKADPLGNLTGTWTAI-----DAGLP-----VCPVTGSLY 141
Db 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGSMSI 211
Qy 142 HLGADKKVKMHESNTAIAIANGLAWSNDLKMYIIDSGRKRV-----EYDYDAST 190
Db 212 IGGIDH-----SLYTGSLWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSI 262
Qy 191 L--SISNOR-PLTFE-----KHEVPGYPDGTIDEENLWVA 225
Db 263 VDSGTTNLRPKKVFEEAAVKSIAASSTEFKPDGFWLGEOLVCWQA 308

RESULT 9

US-09-795-847-51
; Sequence 51, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-795-847-51

Query Match 5.4%; Score 86.5; DB 10; Length 428;

Best Local Similarity 22.0%; Pred. No. 3.2;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

Qy 2 GPVVEKIAEL-----GKY---TVGEGPHDHTQTL-YFVDTVKTF-----HK 41
Db 58 GSFVEMVDNLKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAVGAAPHELHR 111
Qy 42 YVPSQKYYTFCKVKLVSIPIA-----GSPGRFVSLEREIALITWDGVSAAPTSEA 96
Db 112 YQORLSSTYRDLRKG--YVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAAITESDKF 169
Qy 97 IVNVEPHIKNNRLNDGKADPLGNLTGTWTAI-----DAGLP-----VGPVTGSLY 141
Db 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGSMSI 211
Qy 142 HLGADKKVKMHESNTAIAIANGLAWSNDLKMYIIDSGRKRV-----EYDYDAST 190
Db 212 IGGIDH-----SLYTGSLWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSI 262
Qy 191 L--SISNOR-PLTFE-----KHEVPGYPDGTIDEENLWVA 225

Db 263 VDSGTTNLRPKKVFEEAAVKSIAKASTSEKFPDGFNLGEQLVCWQA 308

RESULT 10

US-09-794-743-51
 ; Sequence 51, Application US/09794743
 ; Patent No. US20010021391A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gurney, Mark E.
 ; APPLICANT: Bienkowski, Michael J.
 ; APPLICANT: Heinrikson, Robert L.
 ; APPLICANT: Parodi, Luis A.
 ; APPLICANT: Yan, Riqiang
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
 ; TITLE OF INVENTION: US98
 ; FILE REFERENCE: 28341/6280BC
 ; CURRENT APPLICATION NUMBER: US/09/794,743
 ; CURRENT FILING DATE: 2001-02-27
 ; PRIOR APPLICATION NUMBER: 09/416,901
 ; PRIOR FILING DATE: 1999-10-13
 ; PRIOR APPLICATION NUMBER: 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 51
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
 ; OTHER INFORMATION: delta TM
 US-09-794-743-51

Query Match 5.4%; Score 86.5; DB 10; Length 428;

Best Local Similarity 22.0%; Pred. No. 3.2;
 Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVERIAEL-----GKY---TVGEGPHWDHETQTL-YFVDVVEKTF-----HK 41
 Db 58 GSFVEMVDNLRKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAYGAAPHPLHR 111
 QY 42 YVPSQKKYTFCKVDKLVFIPLA-----GSPGRFVVSLEREITAILTWDGVSAAPTSIEA 96
 Db 112 YYQRLSSTYRDLRKGV--YVPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITESDKF 169
 QY 97 INVVEPHIKNNRLNDGKADPLGNLWCTMAI-----DAGLP-----VGPVTGSLY 141
 Db 170 FIN-----GSNWEGILGLAVAEIARLCGAGFPLNQSEVLASVGSGMI 211
 QY 142 HLGADKKVKMHESNIAIANGLANSLDKMYIIDSGRKRYD-----EVDYDAST 190
 Db 212 IGGIDH-----SLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
 QY 191 L--SISNOR-PLTFE-----KHEVPGYPGQTTIDEGNLMWA 225
 Db 263 VDSGTTNLRPKKVFEEAAVKSIAKASTSEKFPDGFNLGEQLVCWQA 308

RESULT 11

US-09-794-748-51
 ; Sequence 51, Application US/09794748
 ; Patent No. US20020037315A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gurney, Mark E.
 ; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.
 ; APPLICANT: Parodi, Luis A.
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
 ; TITLE OF INVENTION: US98
 ; FILE REFERENCE: 28341/6280JL
 ; CURRENT APPLICATION NUMBER: US/09/794,748
 ; CURRENT FILING DATE: 2001-02-27
 ; PRIOR APPLICATION NUMBER: 09/416,901
 ; PRIOR FILING DATE: 1999-10-13
 ; PRIOR APPLICATION NUMBER: 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 51
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
 ; OTHER INFORMATION: delta TM
 US-09-794-748-51

Query Match 5.4%; Score 86.5; DB 10; Length 428;

Best Local Similarity 22.0%; Pred. No. 3.2;
 Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVERIAEL-----GKY---TVGEGPHWDHETQTL-YFVDVVEKTF-----HK 41
 Db 58 GSFVEMVDNLRKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAYGAAPHPLHR 111
 QY 42 YVPSQKKYTFCKVDKLVFIPLA-----GSPGRFVVSLEREITAILTWDGVSAAPTSIEA 96
 Db 112 YYQRLSSTYRDLRKGV--YVPYTOGKWEGLGTLVSIHPGPNVTVRANIAAITESDKF 169
 QY 97 INVVEPHIKNNRLNDGKADPLGNLWCTMAI-----DAGLP-----VGPVTGSLY 141
 Db 170 FIN-----GSNWEGILGLAVAEIARLCGAGFPLNQSEVLASVGSGMI 211
 QY 142 HLGADKKVKMHESNIAIANGLANSLDKMYIIDSGRKRYD-----EVDYDAST 190
 Db 212 IGGIDH-----SLYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
 QY 191 L--SISNOR-PLTFE-----KHEVPGYPGQTTIDEGNLMWA 225
 Db 263 VDSGTTNLRPKKVFEEAAVKSIAKASTSEKFPDGFNLGEQLVCWQA 308

RESULT 12

US-09-794-925-51
 ; Sequence 51, Application US/09794925
 ; Patent No. US20020064819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gurney, Mark E.
 ; APPLICANT: Bienkowski, Michael J.
 ; APPLICANT: Heinrikson, Robert L.
 ; APPLICANT: Parodi, Luis A.
 ; APPLICANT: Yan, Riqiang
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 28341/6280HI
 ; CURRENT APPLICATION NUMBER: US/09/794,925
 ; CURRENT FILING DATE: 2001-02-27
 ; PRIOR APPLICATION NUMBER: 09/416,901
 ; PRIOR FILING DATE: 1999-10-13
 ; PRIOR APPLICATION NUMBER: 60/155,493

```

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-794-925-51

Query Match          5.4%  Score 86.5;  DB 10;  Length 428;
Best Local Similarity 22.0%;  Pred. No. 3.2;
Matches 63;  Conservative 31;  Mismatches 95;  Indels 97;  Gaps 15;

Qy  2  GPVVEKIAEL-----GKY---TVGEGPHWDHETQTL-YFVDVVEKTF-----HK 41
Db  58  GSFVEMVDNLRGKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAYGAAPHPLHR 111
Qy  42  YVPSQKKYTFCKVDKLVFSFIPLA-----GSPGRFVVSLEREITAILTWDGVSAAPTSIEA 96
Db  112  YVQRLSSTYRDLRKGV--YVPYTOGKWEGLGTDLVSIHPGNVTVVRANIAALTESDKF 169
Qy  97  IVNVPEHIKNRNLNDGKADPLGNLWTGTMAI-----DAGLP-----VGPVTGSLY 141
Db  170  FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGSGMI 211
Qy  142  HLGADKKVKMHESNIAIANGLANSLDKMYIDSCKRRVD-----EYDYDAST 190
Db  212  IGGIDH-----SLYTGSLWYTPIRREWYVEVIIVRVEINGQDLKMDCKEYNYDKSI 262
Qy  191  L--SISNOR-PLTFE-----KHEVPCYPDGQOTIDREGNLWVA 225
Db  263  VDSGTTNLRLPKKVFEAAVKSIAASSTKFPDGFNLGELQVLCWQA 308

RESULT 13
US-09-681-442-51
; Sequence 51, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
; FILE REFERENCE: 28341/6280FG
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-681-442-51

Query Match          5.4%  Score 86.5;  DB 12;  Length 428;
Best Local Similarity 22.0%;  Pred. No. 3.2;
Matches 63;  Conservative 31;  Mismatches 95;  Indels 97;  Gaps 15;

Qy  2  GPVVEKIAEL-----GKY---TVGEGPHWDHETQTL-YFVDVVEKTF-----HK 41
Db  58  GSFVEMVDNLRGKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAYGAAPHPLHR 111
Qy  42  YVPSQKKYTFCKVDKLVFSFIPLA-----GSPGRFVVSLEREITAILTWDGVSAAPTSIEA 96
Db  112  YVQRLSSTYRDLRKGV--YVPYTOGKWEGLGTDLVSIHPGNVTVVRANIAALTESDKF 169
Qy  97  IVNVPEHIKNRNLNDGKADPLGNLWTGTMAI-----DAGLP-----VGPVTGSLY 141
Db  170  FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGSGMI 211
Qy  142  HLGADKKVKMHESNIAIANGLANSLDKMYIDSCKRRVD-----EYDYDAST 190
Db  212  IGGIDH-----SLYTGSLWYTPIRREWYVEVIIVRVEINGQDLKMDCKEYNYDKSI 262
Qy  191  L--SISNOR-PLTFE-----KHEVPCYPDGQOTIDREGNLWVA 225
Db  263  VDSGTTNLRLPKKVFEAAVKSIAASSTKFPDGFNLGELQVLCWQA 308

RESULT 14
US-09-869-414-51
; Sequence 51, Application US/09869414
; Publication No. US20030077226A1
; GENERAL INFORMATION:
; APPLICANT: Beinowski et al.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
; FILE REFERENCE: 28341/6280M
; CURRENT FILING DATE: 2001-06-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-869-414-51

Query Match          5.4%  Score 86.5;  DB 12;  Length 428;
Best Local Similarity 22.0%;  Pred. No. 3.2;
Matches 63;  Conservative 31;  Mismatches 95;  Indels 97;  Gaps 15;

Qy  2  GPVVEKIAEL-----GKY---TVGEGPHWDHETQTL-YFVDVVEKTF-----HK 41
Db  58  GSFVEMVDNLRGKSGQGYVEMTVGSP-----QTLNILDVTGSSNFAYGAAPHPLHR 111
Qy  42  YVPSQKKYTFCKVDKLVFSFIPLA-----GSPGRFVVSLEREITAILTWDGVSAAPTSIEA 96
Db  112  YVQRLSSTYRDLRKGV--YVPYTOGKWEGLGTDLVSIHPGNVTVVRANIAALTESDKF 169
Qy  97  IVNVPEHIKNRNLNDGKADPLGNLWTGTMAI-----DAGLP-----VGPVTGSLY 141
Db  170  FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGSGMI 211
Qy  142  HLGADKKVKMHESNIAIANGLANSLDKMYIDSCKRRVD-----EYDYDAST 190
Db  212  IGGIDH-----SLYTGSLWYTPIRREWYVEVIIVRVEINGQDLKMDCKEYNYDKSI 262
Qy  191  L--SISNOR-PLTFE-----KHEVPCYPDGQOTIDREGNLWVA 225
Db  263  VDSGTTNLRLPKKVFEAAVKSIAASSTKFPDGFNLGELQVLCWQA 308

```

Db 112 YQQLSSTYRDLKGV--YVPYQGWEGELGTLVSIPIHGPNTVVRANIAITAESDKF 169
QY 97 INVNEPHKNNRLNDGKADPLNLTGTMAI-----DAGLP-----VGPVTGSLY 141
Db 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211
QY 142 HLGADKKVKMHESNIAIANGLANSLDKMYIDSGRKRD-----EYDYDAST 190
Db 212 IGGIDH-----SLYTGSLWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
QY 191 L--SISNQR-PLTFEE-----KHEVPGYPDQGTIDEENLWVA 225
Db 263 VDSGTTNLRLPKKVFEAAVKSIRAASTKFPDPGFWLGEQLVCWQA 308

RESULT 15
US-09-548-366-51
; Sequence 51, Application US/09548366
; Publication No. US20030104365A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; FILE REFERENCE: 28341/6280A
; CURRENT APPLICATION NUMBER: US/09/548,366
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
; OTHER INFORMATION: delta TM
US-09-548-366-51

Query Match 5.4%; Score 86.5; DB 12; Length 428;
Best Local Similarity 22.0%; Pred. No. 3.2;
Matches 63; Conservative 31; Mismatches 95; Indels 97; Gaps 15;

QY 2 GPVVEKIAEL-----GKY---TVGEGPHWDHETQTL-YFVDIVYKTF-----HK 41
Db 58 GSFVEMVDNLGRSGQGQYVEMTVGSP-----QTLNILDVTGSSNFAVCAAPHPLHR 111
QY 42 YVPSQKKYTKCKVDKLYSFIPLA-----GSPGRFVYSLEREIAILLTWDGVSAAPTSIEA 96
Db 112 YQQLSSTYRDLKGV--YVPYQGWEGELGTLVSIPIHGPNTVVRANIAITAESDKF 169
QY 97 INVNEPHKNNRLNDGKADPLNLTGTMAI-----DAGLP-----VGPVTGSLY 141
Db 170 FIN-----GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211
QY 142 HLGADKKVKMHESNIAIANGLANSLDKMYIDSGRKRD-----EYDYDAST 190
Db 212 IGGIDH-----SLYTGSLWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSI 262
QY 191 L--SISNQR-PLTFEE-----KHEVPGYPDQGTIDEENLWVA 225
Db 263 VDSGTTNLRLPKKVFEAAVKSIRAASTKFPDPGFWLGEQLVCWQA 308

Search completed: August 1, 2003, 14:43:14
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:11:42 ; Search time 63 Seconds
(without alignments)
4497.926 Million cell updates/sec

Title: US-10-089-986-1

Perfect score: 924

Sequence: 1 atggggccagttgttgaataa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	43.4	4.7	7218	1	US-08-232-463-14
2	42.2	4.6	4739	3	US-08-685-871-1
3	35.8	3.9	1114	4	US-09-378-088A-85
4	35.8	3.9	1152	3	US-08-844-188-42
5	35.8	3.9	1241	4	US-09-378-088A-42
6	35.8	3.9	2132	3	US-08-844-188-39
7	35.8	3.9	2132	4	US-09-378-088A-39
8	34.8	3.8	162450	4	US-09-345-882-1
9	33.6	3.6	2169	4	US-09-434-408-3
c 10	33.6	3.6	43360	4	US-09-453-702B-206
c 11	33.6	3.6	45325	4	US-09-453-702B-261
12	33.2	3.6	1052	4	US-08-961-527-244
13	32.8	3.5	1152	3	US-08-844-188-37
14	32.8	3.5	1152	4	US-09-378-088A-37
15	32.8	3.5	2230	3	US-08-844-188-34
16	32.6	3.5	2230	4	US-09-378-088A-34
17	32.6	3.5	1103	4	US-09-378-088A-57
18	32.6	3.5	14507	3	US-08-785-150-1
19	32.6	3.5	14507	4	US-09-660-239-1
20	32.6	3.5	14507	4	US-09-435-377-1
21	32.4	3.5	270	4	US-09-134-001C-1648
22	32.4	3.5	13425	4	US-08-961-527-151
23	32.2	3.5	3312	4	US-09-060-410-1
24	32.2	3.5	7760	4	US-08-961-527-63
25	32	3.5	3667	4	US-09-221-017B-727
26	31.8	3.4	1725	1	US-08-257-073-14
27	31.6	3.4	624	4	US-08-235-836C-35

c 28	31.6	3.4	1033	4	US-08-936-165A-7	Sequence 7, Appli
c 29	31.6	3.4	8378	5	PCT-US91-09055-1	Sequence 1, Appli
c 30	31.6	3.4	10660	2	US-08-267-803B-8	Sequence 8, Appli
c 31	31.6	3.4	10660	4	US-09-041-886-16	Sequence 16, Appl
32	31.4	3.4	426	4	US-09-328-111-304	Sequence 304, App
33	31.4	3.4	834	3	US-08-633-933A-12	Sequence 12, Appl
34	31.4	3.4	834	3	US-08-844-188-12	Sequence 12, Appl
35	31.4	3.4	834	4	US-09-378-088A-12	Sequence 12, Appl
c 36	31.4	3.4	2313	1	US-08-232-538-5	Sequence 5, Appli
c 37	31.4	3.4	2313	4	US-09-427-353-1	Sequence 1, Appli
c 38	31.4	3.4	2352	1	US-08-232-538-17	Sequence 17, Appl
c 39	31.4	3.4	2352	2	US-08-780-164-17	Sequence 17, Appl
c 40	31.4	3.4	2523	4	US-09-051-363-1	Sequence 1, Appli
c 41	31.4	3.4	2651	2	US-08-786-164-5	Sequence 5, Appli
42	31.4	3.4	5319	1	US-08-169-927-1	Sequence 1, Appli
c 43	31.4	3.4	7032	2	US-08-149-097D-24	Sequence 24, Appl
c 44	31.4	3.4	7032	3	US-08-949-386-24	Sequence 24, Appl
c 45	31.4	3.4	7032	3	US-08-450-562-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT99pt-F1s
; US-08-232-463-14

Query Match

4.7%; Score 43.4; DB 1; Length 7218;

[illegible]

RESULT 2

```

US-08-685-871-1
: Sequence 1, Application US/08685871
: Patent No. 6013499
: GENERAL INFORMATION:
: APPLICANT: IWAMATSU, Shuho
: APPLICANT: IWAMATSU, Akihiro
: TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
: NUMBER OF SEQUENCES: 68
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/685,871
: FILING DATE: 24-JUN-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-184102
: FILING DATE: 25-JUN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 7-262553
: FILING DATE: 14-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16887/845
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 4739 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 448..4509
;
US-08-685-871-1

Query Match      4.6%; Score 42.2; DB 3; Length 4739;
Best Local Similarity 47.0%; Pred. No. 0.0076;
Matches 131; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy      464 ACATAGCTATAGCAAAATGGGCTCGCGTGGAGTAATGATTGAAGAAAATGTATTATATG 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1496 ACACGTGTAGCACCGATTGTACCCGATTAAAGTAGTAGTACATTGATTAATTTTGTATG 1555

Qy      524 ATTCCGGGGAAGAGAGTAGACAGAGTACGATTATGATGCTTCTACATTTATCCATCAGCA 583
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1556 ACATTGGAAGAGATTAAGAGGAGGAGGAACAACATTTCCCTNTTCTTAAGCTTTCGTTGGCA 1615

Qy      584 ATCAACGGCCATTATTTTACTTTTGAAAGCATGAAGTGCCTTGGATATCCAGATGGTCAAA 643
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1616 ATCAACTACCTTTTGTAGGATTTACATATTTATAGCAATCGTAGATACTTATCTTCACGCAA 1675

Qy      644 CAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTTCCAAAGGACACGGAATTAATTA 703
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1676 ATCCTAATGATACAGAACTAGCTCCCAATGCAGATAAAGCTTCACGGAAGACTTTGC 1735

Qy      704 TCAGTACCCCAACCAACCGGAAGTGTTACTGGATACCGTAA 742
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1736 AACCAATCTATAAGCTCTGGAGAACAGCTGCATAATGAAA 1774
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

```

US-09-378-088A-85
; Sequence 85, Application US/09/378088A
; Patent No. 6372480
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; APPLICANT: Michaels, Tracy E.
; APPLICANT: Finstad Lee, Stacy
; APPLICANT: Burmeister, Paula
; APPLICANT: Dojillo, Joanna
; TITLE OF INVENTION: Pesticidal Proteins
; FILE REFERENCE: MAY703C2
; CURRENT APPLICATION NUMBER: US/09/378,088A
; CURRENT FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: US 08/633,993
; PRIOR FILING DATE: 1996-04-19
; PRIOR APPLICATION NUMBER: US 08/844,188
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-378-088A-85

```

Query Match	3.9%	Score 35.8;	DB 4;	Length 1114;
Best Local Similarity	45.5%;	Prod. No. 0.34;		
Matches 127;	Conservative	0;	Mismatches 152;	Indels 0;
Gaps	0			
447	AAAAATGCAGGAGACACATAGCTATAGCAAAATGGGTCGGCTGGAGTAATGATTGAA	506		
618	AAAAATCAATTTGSCACAGCAGCAGTAGGAGTAATGTAGCTTTTACGTCCACATGAAA	677		

Db 705 AAAATATCAATATTGGCAACAGCAGTAGGAGTAATGTAGCTTACGTCACATGAANA 764
QY 507 GAAATGTATTATATTGATTGGGGGAAAGAGTAGACGAGTACGATTATGATGCTTC 566
Db 765 AAAATCATATCTTATGAATGGGACACAGAAATAGATCAAAAACAAATTAATAATAC 824
QY 567 TACATATCCATCAGCAATCAAGCCCATTTATTACTTTTGAAGCATGAAGTGCCTGG 626
Db 825 ATTAGGATTTCAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAATAGGTGG 884
QY 627 ATATCCAGATCGTCAACAAATTTGATGAGGAGGTAAATTTTGGTTGCCCTTTTCCAAGG 686
Db 885 AGGTACAGATCAATATAAACAACACAACTAAATGAAGATTAATAATAGATATAGTCATGA 944
QY 687 ACAGCGAATTTAATAATCAGTACCCCAACACCGGAAGT 725
Db 945 AACTAAATATGGAATAATATCAAGAACAACTCTGAAAT 983

RESULT 6

US-08-844-188-39
; Sequence 39, Application US/08844188
; Patent No. 6127180

GENERAL INFORMATION:

; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Pesticidal Toxins
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,188
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/633,993
; FILING DATE: 19-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-703C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-844-188-39

Query Match 3.9%; Score 35.8; DB 3; Length 2132;
Best Local Similarity 45.5%; Pred. No. 0.45;
Matches 127; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 447 AAAAAATGCACGAGACACATAGCTATTAGCAAAATGGCTCGCGTGGAGTAATGATTTGAA 506
Db 1401 AAAATATCAATATTGGCAACAGCAGTAGGAGTAATGTAGCTTACGTCACATGAANA 1460
QY 507 GAAATGTATTATATTGATTGGGGGAAAGAGTAGACGAGTACGATTATGATGCTTC 566
Db 1461 AAAATCATATCTTATGAATGGGACACAGAAATAGATCAAAAACAAATTAATAATAC 1520
QY 567 TACATATCCATCAGCAATCAAGCCCATTTATTACTTTTGAAGCATGAAGTGCCTGG 626
Db 1521 ATTAGGATTTCAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAATAGGTGG 1580
QY 627 ATATCCAGATCGTCAACAAATTTGATGAGGAGGTAAATTTTGGTTGCCCTTTTCCAAGG 686
Db 1581 AGGTACAGATGAAATATAAACAACACAACTAAATGAAGATTAATAATAGATATAGTCATGA 1640
QY 687 ACAGCGAATTTAATAATCAGTACCCCAACACCGGAAGT 725
Db 1641 AACTAAATATGGAATAATATCAAGAACAACTCTGAAAT 1679

RESULT 7

US-09-378-088A-39
; Sequence 39, Application US/09378088A
; Patent No. 6372480

GENERAL INFORMATION:

; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; APPLICANT: Michaels, Tracy E.
; APPLICANT: Finstad Lee, Stacy
; APPLICANT: Burmeister, Paula
; APPLICANT: Dojillo, Joanna
; TITLE OF INVENTION: Pesticidal Proteins
; FILE REFERENCE: MA703C2
; CURRENT APPLICATION NUMBER: US/09/378,088A
; CURRENT FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: US 08/633,993
; PRIOR FILING DATE: 1996-04-19
; PRIOR APPLICATION NUMBER: US 08/844,188
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-378-088A-39

Query Match 3.9%; Score 35.8; DB 4; Length 2132;
Best Local Similarity 45.5%; Pred. No. 0.45;
Matches 127; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 447 AAAAAATGCACGAGACACATAGCTATTAGCAAAATGGCTCGCGTGGAGTAATGATTTGAA 506
Db 1401 AAAATATCAATATTGGCAACGAGCAGTAGGAGTAATGTAGCTTACGTCACATGAANA 1460
QY 507 GAAATGTATTATATTGATTGGGGGAAAGAGTAGACGAGTACGATTATGATGCTTC 566
Db 1461 AAAATCATATCTTATGAATGGGACACAGAAATAGATCAAAAACAAATTAATAATAC 1520
QY 567 TACATATCCATCAGCAATCAAGCCCATTTATTACTTTTGAAGCATGAAGTGCCTGG 626
Db 1521 ATTAGGATTTCAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAATAGGTGG 1580
QY 627 ATATCCAGATCGTCAACAAATTTGATGAGGAGGTAAATTTTGGTTGCCCTTTTCCAAGG 686
Db 1581 AGGTACAGATGAAATATAAACAACACAACTAAATGAAGATTAATAATAGATATAGTCATGA 1640
QY 687 ACAGCGAATTTAATAATCAGTACCCCAACACCGGAAGT 725

```
Db      1641  AACTAAATATGGAATAATATCAAGAACAACTGAAAT 1679
          |  ||| || ||| |  ||||| | ||| |
RESULT 8
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103806
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134134
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134374
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
; FEATURE:
```


RESULT 10

US-09-453-702B-206/c
; Sequence 206, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43360
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-453-702B-206

Query Match 3.6%; Score 33.6; DB 4; Length 43360;
Best Local Similarity 45.7%; Pred. No. 7.1;
Matches 117; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY	455	ACGAGAGCAACATAGCTATAGCAATGGCTCGCGTGGAGTAATGATTTGAAGAAATGT	514
DB	14108	ATGAGGATAACATTTTAAATTTTATTCATCCTGGTGGTTGAACAATGAGAAAAAATATG	14049
QY	515	ATTATATTGATTCGGGGGAAAGAGAGTAGACGAGTACGATTATGATCTTCTACATTAT	574
DB	14048	GTGAACCTTACATCGTCTAAGAATAGAGTAGAAGAACTGTGTACAGTACTATTAT	13989
QY	575	CCATCAGCAATCAACGGCCATTTTACTTTTGAAGAACGAGTACGCTGGATATCCAG	634
DB	13988	GGAATATTCAATTAATGATTCAGTGGTTTCCAAATGGCTAATGGTATTGAATTTAATC	13929
QY	635	ATGGTCAACAAATTCATGAGGAGGTAATTTATGGTTGCCGTTTCCCAAGCACAGCGAA	694
DB	13928	AAACCCCGCATCTGTTTTTAATCTTAATCTATTAGATTTTGAATTAGAAATCTCAGCAAT	13869
QY	695	TTATTAATAATCAGTAC	710
DB	13868	TAACGCAAAAAAATTC	13853

RESULT 11

US-09-453-702B-261/c
; Sequence 261, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 261:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 261:
US-09-453-702B-261

Query Match 3.6%; Score 33.6; DB 4; Length 45325;
Best Local Similarity 45.7%; Pred. No. 7.2;
Matches 117; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY	455	ACGAGAGCAACATAGCTATAGCAATGGCTCGCGTGGAGTAATGATTTGAAGAAATGT	514
DB	15010	ATGAGGATAACATTTTAAATTTTATTCATCCTGGTGGTTGAACAATGAGAAAAAATATG	14951
QY	515	ATTATATTGATTCGGGGGAAAGAGAGTAGACGAGTACGATTATGATCTTCTACATTAT	574
DB	14950	GTGAACCTTACATCGTCTAAGAATAGAGTAGAAGAACTGTGTACAGTACTATTAT	14891
QY	575	CCATCAGCAATCAACGGCCATTTTACTTTTGAAGAACGAGTACGCTGGATATCCAG	634
DB	14890	GGAATATTCAATTAATGATTCAGTGGTTTCCAAATGGCTAATGGTATTGAATTTAATC	14831
QY	635	ATGGTCAACAAATTCATGAGGAGGTAATTTATGGTTGCCGTTTCCCAAGCACAGCGAA	694
DB	14830	AAACCCCGCATCTGTTTTTAATCTTAATCTATTAGATTTTGAATTAGAAATCTCAGCAAT	14771
QY	695	TTATTAATAATCAGTAC	710
DB	14770	TAACGCAAAAAAATTC	14755

RESULT 12
US-08-961-527-244
; Sequence 244, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1052 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-244
Query Match 3.6%; Score 33.2; DB 4; Length 1052;
Best Local Similarity 50.6%; Pred. No. 2;
Matches 80; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 202 CCTGCGCGTTTGTAGTCAGTTTGGAAAGTGAATAGCCATCTTTACATGGGATGGCGTTT 261
DB 583 CTTGGTTTCTTTGGTCTCACTAAGGAAATCAATCTGTACTTTCTCAGTGGAGTTCT 642
QY 262 AGTGTGACCTACAGCATAGAGCTATTGTTAATGTCGACACACACATTAATAATAC 321
DB 643 GATATTACAGTTCTATCATTTCTGCTGGTGTCAAGCTGACCTCTATGAAACCTTAG 702
QY 322 AGACTCAATGATGGCAAGCAGATCCCTTTGGCAATCT 359
DB 703 GACGTTGATGGTATCTTTGCAGCCCACTGCTGATAT 740
RESULT 13
US-08-844-188-37
; Sequence 37, Application US/08844188
; Patent No. 6127180
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Pesticidal Toxins
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,188
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/633,993
; FILING DATE: 19-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-703C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-844-188-37
Query Match 3.5%; Score 32.8; DB 3; Length 1152;
Best Local Similarity 44.9%; Pred. No. 2.7;
Matches 124; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 447 AAAATGCAGGAGCAACATAGCTATAGCAAAATGGCTCGCTGGAGTAAATGATTGAA 506
DB 618 AAAATATCAATATTGGCAACAGCAGTAGGAAGTAAATGAGTCCGCATGAAA 677
QY 507 GAAATATGATATATTCATTCGGGAAAAGAGTAGACGAGTACGATGATGCTTC 566
DB 678 AAAATCATATGCTTATGAGTGGGTACAGAAATAGATCAAAAACAACTATCATTAATAC 737
QY 567 TACATTATCCATCAGCAATCAACGGCCATTTATTTACTTTTGGAAAGCATGAAGTCCCTGG 626
DB 738 ATTAGGATTTTCAGATTAATATAGATTCCGGAATGAAATTTGATATACCAAGATAGGTGG 797
QY 627 ATATCCAGATGGTCAACAAATTTGATGAGGAGGTAAATTTATGGTTCGCCGTTTTCCAAGG 686
DB 798 AGGTACAGATGAATAAAACACAAATTAACGGAAGAAATTAATAATAGATATAGCCCTGA 857
QY 687 ACAGCGAATTTAAATTCAGTACCCCAACACCGGA 722
DB 858 AACCAAAATAATGGAATAATATCAGAACCAATCAGA 893

RESULT 14
US-09-378-088A-37
; Sequence 37, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; APPLICANT: Michaels, Tracy E.
; APPLICANT: Finstad Lee, Stacy
; US-09-378-088A-37
Query Match 3.5%; Score 32.8; DB 3; Length 1152;
Best Local Similarity 44.9%; Pred. No. 2.7;
Matches 124; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 447 AAAATGCAGGAGCAACATAGCTATAGCAAAATGGCTCGCTGGAGTAAATGATTGAA 506
DB 618 AAAATATCAATATTGGCAACAGCAGTAGGAAGTAAATGAGTCCGCATGAAA 677
QY 507 GAAATATGATATATTCATTCGGGAAAAGAGTAGACGAGTACGATGATGCTTC 566
DB 678 AAAATCATATGCTTATGAGTGGGTACAGAAATAGATCAAAAACAACTATCATTAATAC 737
QY 567 TACATTATCCATCAGCAATCAACGGCCATTTATTTACTTTTGGAAAGCATGAAGTCCCTGG 626
DB 738 ATTAGGATTTTCAGATTAATATAGATTCCGGAATGAAATTTGATATACCAAGATAGGTGG 797
QY 627 ATATCCAGATGGTCAACAAATTTGATGAGGAGGTAAATTTATGGTTCGCCGTTTTCCAAGG 686
DB 798 AGGTACAGATGAATAAAACACAAATTAACGGAAGAAATTAATAATAGATATAGCCCTGA 857
QY 687 ACAGCGAATTTAAATTCAGTACCCCAACACCGGA 722
DB 858 AACCAAAATAATGGAATAATATCAGAACCAATCAGA 893

APPLICANT: Burmeister, Paula
APPLICANT: Dojillo, Joanna
TITLE OF INVENTION: Pesticidal Proteins
FILE REFERENCE: MA703C2
CURRENT APPLICATION NUMBER: US/09/378,088A
CURRENT FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: US 08/633,993
PRIOR FILING DATE: 1996-04-19
PRIOR APPLICATION NUMBER: US 08/844,188
PRIOR FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 130
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 1152
TYPE: DNA
ORGANISM: Bacillus thuringiensis
US-09-378-088A-37

Query Match 3.5%; Score 32.8; DB 4; Length 1152;
Best Local Similarity 44.9%; Pred. No. 2.7;
Matches 124; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 447 AAAATGACGAGAGCAACATAGCTATAGCAAAATGGGCTCGGTGGAGTAATGATTGAA 506
DB 618 AAAATATCAATATTGGCAACAGCAGTAGGAAGTAATGTAGCTTACGTCCGCATGAAA 677
QY 507 GAAATGATATTATTTGATTCGGGGAAGAGAGTAGACGAGTACGATTATGATGCTTC 566
DB 678 AAAATCATATCTTATGATGGGTACAGAAATAGATCAAAAACACTATCATTAATAC 737
QY 567 TACATTATCCATCAGCATCAACGCCATTATTTACTTTTGAAGAAGCATGAAGTCCCTGG 626
DB 738 ATTAGGATTCAGATTAATATAGATTCGGGAATGAATTTGATATACCAAGTAGGTGG 797
QY 627 ATATCCAGATGGTCAAAATTTGATGAGGAGGTAAATTTATGGTTCGGTTTCCCAAGG 686
DB 798 AGGTACAGATGAATATAACACAAATTAACGAAGTAATAAATAGATATAGCCGTGA 857
QY 687 ACAGGGAATTTAAATTCAGTACCCCAACACCGGA 722
DB 858 AACCAAAATATGGAATAATATACGAACATCAGA 893

RESULT 15

US-08-844-188-34
Sequence 34, Application US/08844188
Patent No. 6127180
GENERAL INFORMATION:
APPLICANT: Narva, Kenneth E.
APPLICANT: Schepff, H. Ernest
APPLICANT: Knuth, Mark
APPLICANT: Pollard, Michael R.
APPLICANT: Cardineau, Guy
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Pesticidal Toxins
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,188
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/633,993
FILING DATE: 19-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-703C1
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 2230 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-844-188-34

Query Match 3.5%; Score 32.8; DB 3; Length 2230;
Best Local Similarity 44.9%; Pred. No. 3.6;
Matches 124; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 447 AAAATGACGAGAGCAACATAGCTATAGCAAAATGGGCTCGGTGGAGTAATGATTGAA 506
DB 1575 AAAATATCAATATTGGCAACAGCAGTAGGAAGTAATGTAGCTTACGTCCGCATGAAA 1634
QY 507 GAAATGATATTATTTGATTCGGGGAAGAGAGTAGACGAGTACGATTATGATGCTTC 566
DB 1635 AAAATCATATCTTATGATGGGTACAGAAATAGATCAAAAACACTATCATTAATAC 1694
QY 567 TACATTATCCATCAGCAATCAACGCCATTATTTACTTTTGAAGAAGCATGAAGTCCCTGG 626
DB 1695 ATTAGGATTCAGATTAATATAGATTCGGGAATGAATTTGATATACCAAGTAGGTGG 1754
QY 627 ATATCCAGATGGTCAAAATTTGATGAGGAGGTAAATTTATGGGTTGCCGTTTCCCAAGG 686
DB 1755 AGGTACAGATGAATATAACACAAATTAACGAAGTAATTAATAATAGAAATATAGCCGTGA 1814
QY 687 ACAGGGAATTTAAATTCAGTACCCCAACACCGGA 722
DB 1815 AACCAAAATATGGAATAATATCAGGAACATCAGA 1850

Search completed: August 1, 2003, 13:22:13
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:07:04 ; Search time 1751 Seconds
(without alignments)
8546.333 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgtgaaaa.....cgggagtgaaagtaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.6	11.5	638	13	BI514926
2	98.2	10.6	666	10	AV401529
3	71.2	7.7	670	13	BI611823
4	66.4	7.2	635	13	BI617530
5	65.4	7.1	719	9	AJ425659
6	65.2	7.1	655	13	BI617627

7	65	7.0	560	14	C92366	C92366 C92366 Dict
8	65	7.0	606	14	C92731	C92731 C92731 Dict
9	65	7.0	639	14	C92540	C92540 C92540 Dict
10	65	7.0	679	14	C92417	C92417 C92417 Dict
11	65	7.0	730	14	C90258	C90258 C90258 Dict
12	65	7.0	743	13	BJ435579	BJ435579 BJ435579
13	64.4	7.0	555	9	AI663645	AI663645 u199h04.y
14	63.4	6.9	730	10	AV890554	AV890554 AV890554
15	62.4	6.8	513	13	BJ371993	BJ371993 BJ371993
16	62.4	6.8	582	13	BJ402499	BJ402499 BJ402499
17	62	6.6	592	13	BI161908	BI161908 RE01371.5
18	61.4	6.6	742	13	BJ535587	BJ535587 BJ535587
19	61.2	6.6	704	10	AV905180	AV905180 AV905180
20	61	6.6	583	13	BI364230	BI364230 RE49294.5
21	60.2	6.5	735	13	BJ022574	BJ022574 BJ022574
22	59.6	6.5	524	10	AV839733	AV839733 AV839733
23	59	6.4	457	13	BI592225	BI592225 RH09865.5
24	59	6.4	555	13	BI241345	BI241345 RE38383.5
25	58.6	6.3	515	12	BG639989	BG639989 SD09890.3
26	57.4	6.2	455	13	BI357410	BI357410 RE43884.5
27	57.4	6.2	478	13	BI615830	BI615830 RH45670.5
28	57.4	6.2	480	13	BI361794	BI361794 RE46157.5
29	57.4	6.2	529	13	BI584695	BI584695 RH23952.5
30	57.4	6.2	531	13	BI364664	BI364664 RE49831.5
31	57.4	6.2	546	13	BI367889	BI367889 RE53758.5
32	57.4	6.2	550	13	BI567910	BI567910 RH39220.5
33	57.4	6.2	551	13	BI586938	BI586938 RH27167.5
34	56.6	6.1	415	9	AA387937	AA387937 vc87h05.r
35	56.2	6.1	617	13	BJ522043	BJ522043 BJ522043
36	56.2	6.1	744	13	BI328880	BI328880 602984879
37	55.6	6.1	613	13	BJ373089	BJ373089 BJ373089
38	55.4	6.0	572	13	BJ488689	BJ488689 BJ488689
39	54.8	5.9	491	13	BI365812	BI365812 RE51194.5
40	54.8	5.9	718	13	BJ509252	BJ509252 BJ509252
41	54.6	5.9	590	12	BG638003	BG638003 LD12015.3
42	54.4	5.9	508	9	AU242394	AU242394 AU242394
43	53.4	5.8	593	12	BF607009	BF607009 MY2-00006
44	52.6	5.7	774	9	AI047491	AI047491 ud66dl1.y
45	52	5.6	440	9	AI785853	AI785853 uj56a07.y

ALIGNMENTS

RESULT 1
BI514926
LOCUS BI514926 638 bp mRNA . linear EST 08-APR-2002
DEFINITION BI60016B20D02.5 Bee Brain Normalized Library, BB16 Apis mellifera
CDNA clone BI60016B20D02 5', mRNA sequence.
ACCESSION BI514926
VERSION BI514926.1 GI:15365300
KEYWORDS EST.
SOURCE honeybee.
ORGANISM Apis mellifera

REFERENCE
AUTHORS 1 (bases 1 to 638)
Whitfield, C.W., Band, M.R., Bonaldo, M.F., Kumar, C.G., Liu, L.,
Pardinas, J., Robertson, H.M., Soares, B. and Robinson, G.E.,
TITLE Annotated expressed sequence tags and cDNA microarrays for studies
of brain and behavior in the honey bee
JOURNAL Genome Res. 12 (4), 555-566 (2002)
MEDLINE 21929762
COMMENT Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499
Email: generobi@life.uiuc.edu
This research was funded by the University of Illinois Critical
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation


```

QY 340 GCAGATCCCTTGGCAATCTATGACAGGATACAAATGGCTATTTCACCGTGTCTC 391
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 510 GCAGATCCCTGTTGGAGGTGTTTGTGGAACTATGGTGCATGAAGATCCTC 561

RESULT 3
BI611823
LOCUS
DEFINITION
  RH18245.5prime RH Drosophila melanogaster normalized Head pFlc-1
  Drosophila melanogaster cDNA clone RH18245 5 similar to regucalcin:
  FBan0001803 GO:[ ] located on: X 11A3-11A4:: 08/17/2001, mRNA
  sequence.
ACCESSION
  BI611823
VERSION
  BI611823.1 GI:15507348
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 670)
  Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
  ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
  ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
  Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
  Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
  ,G.M.
  BDGP/HHMI RH Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003706: arm:3R [10470528,10725488]
  estimated-cyto:88D2-88E1: 08/17/2001
  Plate: RH.182 row: D column: 9
  High quality sequence stop: 511.
  Location/Qualifiers
    1..670
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="RH18245"
      /clone_lib="RH Drosophila melanogaster normalized Head
      pFlc-1"
      /sex="male and female"
      /dev_stage="Adult"
      /lab_host="DH5-alpha Tona"
      /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
      BamHI; Library was kindly generated by Piero Carninci at
      the RIKEN. The library was normalized and excised using
      Cre recombinase. Plasmid cDNA library."
BASE COUNT 166 a 176 c 183 g 145 t
ORIGIN
  Query Match 7.7%; Score 71.2; DB 13; Length 670;
  Best Local Similarity 47.9%; Pred. No. 4.5e-09;
  Matches 273; Conservative 0; Mismatches 288; Indels 9; Gaps 2;

QY 47 TTGAGAGAGTCTCTACTGGATCATGAACATCAGACCTTATATTTTCGTGCACACCGTAG 106
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 93 TGGCGAGGAGACCCCATGGGTGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 152
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 107 AGAAAACCTTTTCATAATATGTACCTTCTCAGAAAAAATACACGTTTGTGTA---G 160
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 153 CCGCGCGCATTAATCGTATGATTTCAGCAGAACAAAGTGACAGGCTAAATTCGAGG 212
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 161 TAGATAAATCGTTTCTTCATTATTCCTTCCTGATCCCTGGCCGTTTGTAGTCA 220
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 213 GCGAGATATTGTCATCGTTCTATTCGCGGTGAGACAAACCCGACGAGTTTGGCCGTAG 272
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 221 GTTTGGAAACGTGAATAGCCATTCTTTACATGGGATGGCGGTAGTGTGCTGCACCTACAAGCA 280
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 273 GATCGGTCTTCGTACGGTCAATCGTCCAGTGGATGGAGTCTCCGAGTGGCCAAAGTCA 332
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 281 TAGAAGCTATTGTAATGTCGAACCAATTAATAAATAACAGACTCAATGATGGCAAG 340
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 333 CTCGCACCCCTGTTTCGAGGTCCACCGGACCTGAAGGAACCCGCTTAATGATGCCAANA 392
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 341 CAGATCCCTTGGCAATCTATGACAGGATACAATGGCTATTTAGCGTGTCTCCCGTAG 400
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 393 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGGCATATTCA 449
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 401 GACCGGTCACTGCGAGTTTATATCATTTAGGGGCTGATAAAAGGTAAAAATGCACGAGA 460
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 450 CCAATGGAAGGTGAGCTCTACAGCTGGCAGCGCGTGGACGCCCAACGCTATCCGTA 509
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 461 GCAACATAGCTATAGCAAAATGGCTGCGTGGAGTAAATGATTTGAAGAAAATGATTATA 520
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 510 GCAAGTGGGCATATCCATGGCTGGCTGGATGTCAAGGCCAAGAAGTTCTACTTCA 569
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 521 TTGATTGGGGAAAGAGTAGACGAGTACGATTATGATGCTTTCATATTTATCCATCA 580
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 570 TCGACACCAACACACGAGGTATTTGGCTTATGACTACCATCAGACCGCGCGCTAA 629
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 581 GCAATCAACGCCCATTTATTTACTTTTGAAA 610
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 630 GCAACCCAAAGTCATTTTCGATCTGAGGAA 659
||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
BI617530
LOCUS
DEFINITION
  RH47894.5prime RH Drosophila melanogaster normalized Head pFlc-1
  Drosophila melanogaster cDNA clone RH47894 5 similar to regucalcin:
  FBan0001803 GO:[ ] located on: X 11A3-11A4:: 08/18/2001, mRNA
  sequence.
ACCESSION
  BI617530
VERSION
  BI617530.1 GI:15513055
KEYWORDS
  EST.
SOURCE
  fruit fly.
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 635)
  Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
  ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
  ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
  Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
  Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
  ,G.M.
  BDGP/HHMI RH Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AE003706: arm:3R [10470528,10725488]
  estimated-cyto:88D2-88E1: 08/18/2001
  Plate: RH.478 row: H column: 10
  High quality sequence stop: 528.
  Location/Qualifiers
    1..635
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="RH47894"
      /clone_lib="RH Drosophila melanogaster normalized Head
      pFlc-1"
      /sex="male and female"
      /dev_stage="Adult"
      /lab_host="DH5-alpha Tona"
      /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:

```


BASE COUNT	164 a	171 c	178 g	142 t	ORIGIN
Query Match	7.1%	Score 65.2	DB 13	Length 655;	
Best Local Similarity	47.6%	Pred. No. 2.2e-07			
Matches 261; Conservative	0	Mismatches 278	Indels 9	Gaps 2;	
QY	47	TTGGAGAAGTCTCTACTGGGATCATGAACCTACAGACCTTATATTTGTCGCACACCGTAG	106		
Db	93	TGGCGAGGACCCCACTGGGATGTTGATCGCCAGAGCTGTACTACGTGGACCTCGAAT	155		
QY	107	AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACAGTTTGTGAAA-----G	160		
Db	153	CGCCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAAATCGAGG	212		
QY	161	TAGATAAATCGTGTCTTTTCATTATTCCTCTGCTGGATCCCTCGGCGGTTTGTACTGCA	220		
Db	213	GCAGATATTTGCATCGTTCATTTCTGCCGTTTGAGAACAAACCGCAGGAGTTTGCCGTAG	272		
QY	221	GTTTGGAACGTGAATAGCAATCTTACATGGGATGGCGTTAGTGTCTGCACCTACAAGCA	280		
Db	273	GATGCGGTCTTCGTACGGTCAATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAAGTCA	332		
QY	281	TAGAAGCTATTGTTAATGTGCAACCAACACATTTAAAAATAACAGACTCAATGTATGGCAAG	340		
Db	333	CTGCACCCCTGTTTCGAGGTGCAGCCGACCTGAAGAAACACCGCCTTAATGATGCCAAAA	392		
QY	341	CAGATCCCTTGGCAATCTATGACACGGTACATGGCTATTGACGTGGTGTCTCCCCGTAG	400		
Db	393	CCGATCCCAATGGCCGCTTTTACGGTGGCACCATTGGC---CGACAGTGGCGACATATTCA	449		
QY	401	GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGTAAAAATGCACGAGA	460		
Db	450	CCCAATGGAAGGTGAGCTCTACAGTGGCAGGCCGGTGGACGCCCAACGCTATCCGTA	509		
QY	461	GCAACATAGCTATAGCAAAATGGGCTCGCGTGGAGTATATGATTTTGAAGAAAATGTTATTAT	520		

Qy 712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCTCAGGTCACTCTGTA 771

```

Db      331 ATCACTAAAAAATTTTATTAACTGTAACCTCTTCCAAATGCTCTCAAGGTCACCTCTTGT 390
QY      772 GCATTGGCGTCCGAATTTGGATGAACATGCAATCTGCTGGTCTTCAGCTTGAC 831
Db      391 GCTTTTGGTGTCTGATCTCAATCTCTTTATATCAACAAGCAAGCGTGATGTTGAA 450
QY      832 GACAGTTCT 840
Db      451 COTGATTCT 459

RESULT 8
LOCUS   C92731 606 bp mRNA linear EST 12-JUL-1999
DEFINITION C92731 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSF142, mRNA sequence.
ACCESSION C92731.1 GI:3074607
VERSION C92731.1
KEYWORDS EST
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
REFERENCE 1 (bases 1 to 606)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source Location/Qualifiers
1. .606
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSF142"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 207 a 86 c 119 g 194 t
ORIGIN

Query Match 7.0%; Score 65; DB 14; Length 606;
Best Local Similarity 50.4%; Pred. No. 2.5e-07;
Matches 186; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 472 ATAGCAATGGCTCGCGTGGAGTAATGATTTGAAGAAAATGATTAATATTGATCGGGG 531
Db 141 ATTGGTAATGGTGGTGGTATGCAATTTGATGATCAAAAAGTTTACTATATATGATACACCA 200
QY 532 AAGAAGATGACGAGTACGATGATGATGCTTCTACATTTCCATCAGCAATCAACGG 591
Db 201 ACTTTGAATATCGATGAATGGAATTAACGCTGCGCCCAAGGCACTGTATCAATTCGCT 260
QY 592 CCATTTATTTATTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAAACAATTGAT 651
Db 261 CATTTGATATCTTTGAAAACCGTAATTC---GGCTTCCAGATGGTATGACCATTTGAT 317
QY 652 GAGAGGGTAATTTATCGGTTGCGTTTCCAGAGGACAGCAATTTAAATCACTACC 711
Db 318 AGTGAAGGTAACCTCTGGATGCTCATTTGGGAGGGTGGTAGAGTTACAAGATGGTGCCA 377
QY 712 CAACAACCGGAAGTGTACTGGATACCGTAAAAAATACCAGATCCTCAGGTCACTCTGTA 771

```

```

Db      378 ATCACTAAAAAATTTTATTAACTGTAACCTCTTCCAAATGCTCTCAAGGTCACCTCTTGT 437
QY      772 GCATTGGCGTCCGAATTTGGATGAACATGCAATCTGCTGGTCTTCAGCTTGAC 831
Db      438 GCTTTTGGTGTGATCTGATCTTCAATCTCTTTATATCAACAAGCAAGCGTGATGTTGAA 497
QY      832 GACAGTTCT 840
Db      498 COTGATTCT 506

RESULT 9
LOCUS   C92540 639 bp mRNA linear EST 12-JUL-1999
DEFINITION C92540 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSE589, mRNA sequence.
ACCESSION C92540
VERSION C92540.1 GI:3074416
KEYWORDS EST
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
REFERENCE 1 (bases 1 to 639)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source Location/Qualifiers
1. .639
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSE589"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 233 a 87 c 120 g 199 t
ORIGIN

Query Match 7.0%; Score 65; DB 14; Length 639;
Best Local Similarity 50.4%; Pred. No. 2.5e-07;
Matches 186; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 472 ATAGCAATGGCTCGCGTGGAGTAATGATTTGAAGAAAATGATTAATATTGATCGGGG 531
Db 150 ATTTGTAATGGTGGTGGTATGCTCAATTTGATCAAAAAGTTTACTATATTTGATACACCA 209
QY 532 AAAAGAAGATGACGAGTACGATGATGATGCTTCTACATTTCCATCAGCAATCAACGG 591
Db 210 ACTTTGAATATCGATGAATGGAATTAACACGCTGCCCAAGCACTGTATCAATTCGCT 269
QY 592 CCATTTATTTTACTTTTGAAGAAGCATGAAGTGCCTGGATATCCAGATGGTCAAAACAATTGAT 651
Db 270 CATTTGATATCTTTTGAAGACCGTAATTC---GGCTTCCAGATGGTATGACCATTTGAT 326
QY 652 GAGAGGGTAATTTATGGGTTGCCCTTTTCCAAAGGACAGCAATTTAAATCACTACC 711
Db 327 AGTGAAGGTAACCTCTGGATGCTCATTTGGGAGGGTGGTAGAGTTACAAGATGGTGCCA 386
QY 712 CAACAACCGGAAGTGTACTGGATACCGTAAAAAATACCAGATCCTCAGGTCACTCTCTGTA 771

```


Db	430	ATCACTAAAAACCTTTTAACTGCTAACTGTTCCAAATGCTCAAGGGTCACTTCTTCT	489
Qy	772	GCATTTGGCGGTCCGAAATTTGGATGAAGCTGCATGTAACATCTGCTGGTCTTCAGCTTGGAC	831
Db	490	GTCTTTTGGTATCTGATCTTCAATCTCTTTATATCAACAACGTCAAGACGTGATGGTGAA	549
Qy	832	GACAGTTCT	840
Db	550	CCTGATCT	558

RESULT 11	
C90258	
LOCUS	730 bp mRNA linear EST 20-APR-1999
DEFINITION	C90258 Dictyostelium discoideum SS (H.Urushihara) Dictyostellium discoideum CDNA clone SSI507, mRNA sequence.
ACCESSION	C90258
KEYWORDS	C90258.1 GI:3059878
SOURCE	EST.
ORGANISM	Dictyostelium discoideum.
REFERENCE	Dictyostelium discoideum
AUTHORS	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostellium. 1 (bases 1 to 730)
	Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H., Williams,R., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,I.
TITLE	Unpublished CDNA in Dictyostelium discoideum
JOURNAL	Unpublished (1998)
COMMENT	Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan Tel: 81-298-53-4664 Fax: 81-298-53-6614 Email: hideko@biol.tsukuba.ac.jp.

```

/rd_xref="taxon:44689"
/clone="SSI507"
/clone.lib="Dlctyostelium discoideum SS (H.Urushi-hara)"
/dev_stage="slug"
BASE COUNT      263 a    101 C    135 g    231 t
ORIGIN

Query Match      7.0%; Score 65; DB 14; Length 730;
Best Local Similarity 50.4%; Pred. No. 2.6e-07;
Matches 186; Conservative 0; Mismatches 180; Indels 3; Gaps 1

```

[illegible]

[illegible]

	Query Match	6.9%	Score 63.4	DB 10	Length 730
	Best Local Similarity	52.6%	Pred. No. 7.5e-07		
	Matches 214	Conservative 0	Mismatches 181	Indels 12	Gaps 3
Oy	412	GGCAGTTTATATCATTTTAGGGCTGATAAAAAGTAAAAATGACGAGAGCAACATAGCT	471		
Db	718	GGCAAGCTTATTGTCCTGTGACATTGATCATTCGTTAAAGACAAAGGTTTATCCGGTAGAT	659		

	Query Match	6.8%	Score 62.4;	DB 13;	Length 513;
	Best Local Similarity	50.9%	Pred. No. 1.3e-06;		
	Matches 175;	Conservative 0;	Mismatches 166;	Indels 3;	Gaps 1;
Qy	472 ATACCAATGGGCTCGCGTGGAGTAATGATTTGAGAAATGTATTATATTGATTCGGG	531			
Db	341 ATTCGTAAATGGTGGTATGGTCAATGTATGATCAAAAAAGTTTTACTATATTGATACACCA	282			
Qy	532 AAAAGAAGATAGACGAGTACGATTATGATGCTTCTACATTATTCACATCAGCAATCAACGG	591			
Db	281 ACTTTGAATATCATGAATAATGGATTACACGCTGCCCAAGCACTGATTCAAATTCGTCGT	222			
Qy	592 CCATTATTTACTTTTGAAGACGATGAAGTGCCTGGATATCCAGATGCTCAACAACTTGTAT	651			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:01:21 ; Search time 292 Seconds

(without alignments)
7126.186 Million cell updates/sec

Title: US-10-089-986-1

Perfect score: 924

Sequence: 1 atggggccagttgtgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_l01002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	863	93.4	865	22 AAF86444	Firefly luciferin
2	325.2	35.2	924	24 ABL49712	Luciola lateralis
3	320.4	34.7	930	24 ABL41998	Nucleotide sequenc
4	98.6	10.7	912	23 ABL12833	Drosophila melanog
5	90.4	9.8	1216	16 AAQ87295	Rat regucalcin cDN
6	80.8	8.7	1573	24 ABL99968	Mouse SMP30 encodi
7	74.4	8.1	1356	16 AAQ90035	Human SMP30 gene.
8	74.4	8.1	1356	20 AAX40196	Senescence marker
9	74.4	8.1	1356	24 ABN95150	Gene #1648 used to

C	10	70.4	7.6	2970	23	ABL12832	Drosophila melanog
	11	67	7.3	1086	23	ABL03467	Drosophila melanog
	12	67	7.3	1094	23	ABL19751	Drosophila melanog
	13	67	7.3	1119	23	ABL03469	Drosophila melanog
	14	52.8	5.7	3554	23	ABL03468	Drosophila melanog
	15	52.8	5.7	4624	23	ABL03466	Drosophila melanog
	16	52.8	5.7	6625	23	ABL19750	Drosophila melanog
	17	46.8	5.1	1346	22	AAH57498	Human liver cell s
	18	42.2	4.6	4065	19	AAV42941	Nucleic acid encod
	19	42.2	4.6	4065	20	AAK39676	Renal cancer assoc
	20	42.2	4.6	4065	24	ABL59542	Human ROCK1 cDNA S
	21	42.2	4.6	4739	18	AAW78203	Physiologically ac
	22	41.2	4.5	2365589	24	ABA90521	Genomic sequence o
C	23	40.6	4.4	2641	23	ABL28434	Drosophila melanog
	24	39.8	4.3	428	24	ABO58878	Human colon cancer
	25	39.8	4.3	4590	22	AAH24065	Yeast AOD9604-asso
	26	38.8	4.2	684707	24	ABO67196	Listeria innocua c
	27	38.8	4.2	3011208	24	ABO69245	Listeria innocua D
C	28	38.6	4.2	11812	22	AAK45501	Chemically pretrea
	29	38.6	4.2	11812	22	AAK46741	Tumour suppressor
	30	38.6	4.2	11812	24	ABL34118	Human immune syste
C	31	38.6	4.2	11812	24	ABK28431	DNA transcription
	32	38	4.1	66986	22	AAF28542	Genomic fragment #
C	33	37.8	4.1	1993	24	ABQ70703	Listeria monocytog
	34	36.6	4.0	3078	24	ABN59645	Novel human coding
C	35	35.8	3.9	1114	22	AAF91977	B. thuringiensis 44
	36	35.8	3.9	1152	18	AAV06394	B. thuringiensis 4
	37	35.8	3.9	1152	21	AAA96766	DNA sequence encod
	38	35.8	3.9	1241	22	AAF91952	B. thuringiensis 44
	39	35.8	3.9	1521	22	AAF92018	B. thuringiensis 58
	40	35.8	3.9	2132	18	AAV06392	B. thuringiensis 1
	41	35.8	3.9	2132	21	AAA96764	Nucleotide sequenc
	42	35.8	3.9	2132	22	AAF91950	B. thuringiensis 14
	43	35.2	3.8	32038	22	AAS30028	Human lung antigen
	44	35.2	3.8	32038	22	AAK89980	Human digestive sy
	45	35.2	3.8	32038	22	AAK90877	Human digestive sy

ALIGNMENTS

RESULT 1
AAF86444 ID AAF86444 standard; DNA; 865 BP.
XX AAF86444;
AC AAF86444;
XX AAF86444;
DT 26-JUN-2001 (first entry)
XX AAF86444;
DE Firefly luciferin regenerating protein coding sequence.
XX AAF86444;
KW Firefly; luciferin regenerating protein; oxyluciferin; luciferase; ds.
XX AAF86444;
OS Photinus pyralis.
XX AAF86444;
FH Key Location/Qualifiers
FT CDS 2..865
FT FT /*tag= a
FT FT /product= "Luciferin regenerating protein"
FT FT /transl_except= (pos:782..784, aa:leu)
FT FT /note= "No start or stop codon given"
XX AAF86444;
PN WO200125426-A1.
XX AAF86444;
PD 12-APR-2001.
XX AAF86444;
PF 22-SEP-2000; 2000WO-JP06527.
XX AAF86444;
PR 06-OCT-1999; 99JP-0285258.
XX AAF86444;
PA (KIKK) KIKKOWAN CORP.
XX AAF86444;
PI Kurosawa K, Kajiyama N;

XX	WPI; 2001-266307/27.
XX	P-PSDB; AAB82087.
XX	Luciferin regenerating protein and gene encoding it useful for
PT	regenerating expensive luciferin from oxyluciferin and D-cysteine -
XX	Disclosure; Page 14-15; 21pp; Japanese.
XX	The present sequence is the coding sequence from firefly for a protein
CC	that regenerates luciferin by using oxyluciferin and D-cysteine. The
CC	encoded protein can be used for regeneration of luciferin, a substrate
CC	for luciferase, used for ATP (adenosine triphosphate) assays in both
CC	medical and food hygiene areas.
XX	Sequence 865 BP; 267 A; 163 C; 195 G; 239 T; 1 other;
XX	

Query Match	93.4%	Score 863	DB 22	Length 865
Best Local Similarity	100.0%	Pred. No. 3.1e-242		
Matches	864	Conservative 0	Mismatches 0	Indels 0
			Gaps 0	
Qy	61	CAC TGGGATCATGAACACTCAGACCTTATATTTCTGTCGACACCGTAGAGAAAAC TTTTTCAT	120	
Db	2	CAC TGGGATCATGAACACTCAGACCTTATATTTCTGTCGACACCGTAGAGAAAAC TTTTTCAT	61	
Qy	121	AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAAGTAGATAAAAC TGGTTCTTTTC	180	
Db	62	AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAAGTAGATAAAAC TGGTTCTTTTC	121	
Qy	181	ATTATTTCCCTTGCCTGGATCCCTTGGCCGTTTTGTAGTCAGTTTGGACGCTGAATAGCC	240	
Db	122	ATTATTTCCCTTGCCTGGATCCCTTGGCCGTTTTGTAGTCAGTTTGGACGCTGAATAGCC	181	
Qy	241	ATTCCTTACATGGATGGCGTTAGTGCTGCACCTACAAAGCATAGAAGCTATTGTTAATGTC	300	
Db	182	ATTCCTTACATGGATGGCGTTAGTGCTGCACCTACAAAGCATAGAAGCTATTGTTAATGTC	241	
Qy	301	GAACACACATTTAAAATAACAGACTCAATGATGGCAAGACAGATCCCTTGGCAATCTA	360	
Db	242	GAACACACATTTAAAATAACAGACTCAATGATGGCAAGACAGATCCCTTGGCAATCTA	301	
Qy	361	TGCACAGGTACAAATGGCTATTGACGCTGGTCTCCCGTAGGACCGGTCTACTGGCAGTTTA	420	
Db	302	TGCACAGGTACAAATGGCTATTGACGCTGGTCTCCCGTAGGACCGGTCTACTGGCAGTTTA	361	
Qy	421	TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAT	480	
Db	362	TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAT	421	
Qy	481	GGGCTCGCTGGAGTAATGATTGTAAGAAAAATGTTATATTGATTCGGGGAAGAACGA	540	
Db	422	GGGCTCGCTGGAGTAATGATTGTAAGAAAAATGTTATATTGATTCGGGGAAGAACGA	481	
Qy	541	GTACAGCAGTACGATTATGATGCTTCTACATTTATCCATCAGCAATCAACGCCCATTTAT	600	
Db	482	GTACAGCAGTACGATTATGATGCTTCTACATTTATCCATCAGCAATCAACGCCCATTTAT	541	
Qy	601	ACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGGTCAACAATTTGATGAGGAGGT	660	
Db	542	ACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGGTCAACAATTTGATGAGGAGGT	601	
Qy	661	AAHTTATGGGTTGCCGTTTTTCCAAGGACAGCGAATTTAAATACGTACCCCAACACCG	720	
Db	602	AAHTTATGGGTTGCCGTTTTTCCAAGGACAGCGAATTTAAATACGTACCCCAACACCG	661	
Qy	721	GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGTGCACCTCTGTAGCAATTTGGC	780	
Db	662	GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGTGCACCTCTGTAGCAATTTGGC	721	
Qy	781	GGTCCGAATTTGGATGAAC TGCATGTACATCTGCTGGTCTTTACGCTTGACGACAGTTCT	840	
Db	722	GGTCCGAATTTGGATGAAC TGCATGTACATCTGCTGGTCTTTACGCTTGACGACAGTTCT	781	

Qy	841	TTTNGACAAAGTTT	AGTTT	AGTTT	AAATGGGCACGCT	TACAG	AGT	TACAG	AGTTT	TAGCGCTCAAAGGT	900
Db	782	TTTNGACAAAGTTT	AGTTT	AGTTT	AAATGGGCACGCT	TACAG	AGT	TACAG	AGTTT	TAGCGCTCAAAGGT	841
Qy	901	TTTCGGGGGAGTT	ATAAGT	CAAGCTA	924						
Db	842	TTTCGGGGGAGTT	ATAAGT	CAAGCTA	865						

RESULT 2

ABL9712	
ID	ABL49712 standard; DNA; 924 BP.
XX	
AC	ABL49712;
XX	
DT	29-MAY-2002 (first entry)
XX	
DE	Luciola lateralis luciferin regenerating protein encoding DNA SEQ:1.
XX	
KW	Luciola lateralis; luciferin; regeneration; Japanese firefly; adenosine triphosphate; ATP; medical science; food hygiene; luciferase; luminescence; gene; ds.
XX	
OS	Luciola lateralis.
XX	
FH	Key Location/Qualifiers
FT	1..924
CDS	/tag= "a"
FT	/product= "Luciferin regeneration protein"
FT	
XX	
XX	WO200210384-A1.
PJ	
PD	07-FEB-2002.

WO200210384-A1

07-FEB-2002

26 - TWT - 2001: 2001WQ-TP06455

29-III-2000: 2000TP-0228227

"THEY" - "THEY" MAN CORP

Uchida, K. 1993. *Uchida, K.*

[illegible]

P-PSDB; ABB06339.

Luciola lateralis-originated genes encoding proteins capable of regenerating luciferin especially from oxyluciferin, useful for producing recombinant DNAs and transformants to give proteins useful in assaying adenosine triphosphate -

Claim 9: page 17-18: 23pp: Japanese.

the present sequence encodes a protein capable of regenerating

The present sequence encodes a protein capable of regenerating luciferin, which is isolated from *Luciola lateralis* (Japanese firefly). The gene encoding the protein capable of regenerating luciferin can be used for producing recombinant DNAs and transformants, which can be used for the production of proteins useful in assaying adenosine triphosphate (ATP) in medical sciences and food hygiene. The protein can be added to the luciferin-luciferase reaction system to sustain limiting luciferin and reduce their amounts used.

[illegible]

	20	30	40	50	60	70	80	90	100
Length	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Weight	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Volume	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Area	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Perimeter	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Mass	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Density	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Pressure	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Temperature	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Speed	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Acceleration	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Force	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Momentum	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Energy	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Power	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Frequency	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Wavelength	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Amplitude	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Phase	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Velocity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Angular Velocity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Angular Acceleration	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Torque	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Work	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Heat	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Entropy	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Enthalpy	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Gibbs Free Energy	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Helmholtz Free Energy	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Chemical Potential	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Fugacity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Activity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Concentration	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Partial Pressure	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Vapor Pressure	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Osmotic Pressure	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Bulk Modulus	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Young's Modulus	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Poisson's Ratio	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Shear Modulus	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Thermal Conductivity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Electrical Conductivity	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Thermal Expansion Coefficient	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Compressibility	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Refractive Index	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Dielectric Constant	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Magnetic Susceptibility	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0

very Match
35.2%; Score 325.2; DB 2

	1:	2:	3:
Local Similarity	60.2%	Mismatched 264.	Totals
Global Similarity	8.2e-83;		

60

1 ATGGGGCCAGTTGTTGAAAAAATTCAGAACCTTGGCAAGTATACGGTTGGAGAAAGGCTT

1 TCGGAGGGTCCA
2 CACATTTCCAAATCGGAGGGTCCA 60

I ATGTCGCCAGTATAATGAACAGATCAC TGAAGTAGHCAATATCCAAATCGGAGAGCGGACG

51 CAGTCCCTCTCTCAAACTCAGACCTTATTTTCGTCGACACCCGTAGAGAAACTTTTCAT 12

6T CACTGGGATCATGAAGACCTCAAGACCTTAATAATTCCCTCGAAGGGTTTCGTGTTTG
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 61 CACTGGGATACCGAACAACAAAGTTTGTATTTTGGATATTTCTAGAAAAATCCATACAC 120
 Qy 121 AATATATGATCTTCTCAGAAAAATACACGTTTTTGTAAAGPAGATAAATCGTTTCTTTC 180
 Db 121 AATATGATCTTCTCAGAAAAATACACGTTTTTGTAAAGPAGATAAATCGTTTCTTTC 180
 Qy 181 ATTATTTCCCTTGTCTGATCCCTGGCGTTTGTAGTCACTTTTGGAGAGCTGAATAGCC 240
 Db 181 ATTATTTCCCTTGTCTGATCCCTGGCGTTTGTAGTCACTTTTGGAGAGCTGAATAGCC 240
 Qy 241 ATTCTTACATGGGATGGCTAGTGTGCTGACCTTACAGCATGAGAGCTATTGTAATGTC 300
 Db 241 GTTCTTACTTGGGATGGCTAGTGTGCTGACCTTACAGCATGAGAGCTATTGTAATGTC 300
 Qy 301 GAACACACATTAATAAATACAGACTCAATGATGGCAAGCAGATCCCTTTGGCAATCTA 360
 Db 301 GATACGGGAATCGAAGGAATACATTCATGATGTAAGCAGATGCGTTTGGCAATTTG 360
 Qy 361 TGGACAGGTACAATGGCTATTGACGCTGCTGCTCCCGTAGGACCGGTCACTGGCAGTTA 420
 Db 361 TGGCAGGTACAATGATGTAATTCGATATTGAAACCAAGGTCCTTAATACAGGAACA 420
 Qy 421 TATCATTTAGGGCTGTATAAAGGTAATAATGCAGGAGCAACATAGCTATTAGCAAT 480
 Db 421 TTGTACAGCTGTCTATACAGCACTTAAGAAACATATTTCTAATATCTTCTTATCAAT 480
 Qy 481 GGGCTCGGTGAGTAATGATTTGAAGAAATGATTATTTATTTGTTGGGAAAGAGA 540
 Db 481 GGCCTAGCTGGAATTAAGACTCAAAAAATTTTATTTATCGACTCCAATAAAGAGA 540
 Qy 541 GTACAGAGTACGATATTGATGCTTTCATATTCCATTCAGCAATCAACGGCCATATT 600
 Db 541 ATAGATCAGTTGATGATGCTTTCGAAATTTAATTTATCAAAATTTGTCACCACTGTT 600
 Qy 601 ACTTTTGAAGAAGTGAAGTCCCTGGATATCCAGATGCTCAACAAATTTGATGAGGAGGT 660
 Db 601 ACTCTGCACAAACATGGGATACAGGTTTACTGATGCCCAACGATAGTGAATATGAT 660
 Qy 661 AATTTATGGTGGCTGTTTCCAAAGGACAGCAATTTAATAATCAGTACCCCAACACCG 720
 Db 661 AATTTATGGTGGCTGTTTCCAAAGGACAGCAATTTAATAATCAGTACCCCAACACCG 720
 Qy 721 GAAGTGTACTGGATACCGTAAATACAGATCCTCAGTCACTCTGTAGCATTTGGC 780
 Db 721 GAATCTTGTGTTGTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Qy 781 GGTCCGAATTTGGATGAACATGCTAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 GGATCAAGTTGGACGAACCTTATGTAACGACTTCTGCTATCAAGAGTATGAA---CT 837
 Qy 841 TTNGACAAAGTTTGTATGATGGCAGCTCTACAGATTAACAGGTTTGGCGTCAAGGT 900
 Db 838 GACTCTACAAACCTAGTAAGGGTGGATTTGTACAGAGTTACTGGATTAGGTGTTAAAGGT 897
 Qy 901 TTCGCGGAGTTAAAGTGAAGCT 923
 Db 998 TTACCCGCACACAGATTCACTCT 920

RESULT 3

ABL41998

ID ABL41998 standard; DNA; 930 BP.

XX ABL41998;

AC ABL41998;

DT 11-JUN-2002 (first entry)

DE Nucleotide sequence of a protein capable of regenerating luciferin.

XX Luciferin regenerating protein; luciferin; oxyluciferin; luminescence;

KW ATP assay; gene; ss.

XX

OS Luciola cruciata..

XX Key Location/Qualifiers
 FH CDS 1..930
 FT /*tag= a
 FT /product= "Luciferin regenerating protein"
 XX WO200210383-A1.
 PN 07-FEB-2002.
 XX 26-JUL-2001; 2001WO-JP06454.
 XX 28-JUL-2000; 2000JP-0228226.
 XX (KIKK) KIKKOWAN CORP.
 XX Hirokawa K, Kurosawa K, Kajiyama N;
 XX WPI; 2002-180080/23.
 XX P-PSDB; ABB09720.
 XX New Luciola cruciata-originated genes encoding proteins capable of regenerating luciferin especially from oxyluciferin, for producing recombinant DNAs and transformants -
 PT Claim 9; Page 17-18; 23pp; Japanese.
 XX The present sequence encodes a protein capable of regenerating luciferin. The protein can especially regenerate luciferin from oxyluciferin. The protein can be added to the luciferin-luciferase reaction system to sustain luminescence and reduce the amounts used. The polynucleotide sequence is useful for producing recombinant transformants, which are useful for the production of proteins useful in assaying ATP in medical sciences and food hygiene.
 XX Sequence 930 BP; 305 A; 150 C; 203 G; 272 T; 0 other;

Query Match 34.7%; Score 320.4; DB 24; Length 930;
 Best Local Similarity 60.6%; Pred. No. 2.1e-83;
 Matches 563; Conservative 0; Mismatches 357; Indels 9; Gaps 2;

Qy 1 ATGGGGCCAGTTTGTAAAAAATTCAGAACTTGGCAAGTATACGGTTGGAGAGGTCCT 60
 Db 1 ATGGCTCCAACTTGTGAACAAATAGTGAATGGGCACCTTATTTGCTTGCAGAAAGTCCA 60
 Qy 61 CACTGGGATCAAACTCAGACCTTATATTTCTGACACCGTAGAGAAACTTTTCAT 120
 Db 61 CATTGGGACGAGAACTCAAAAGTTTGTACTTTCTGGATATTGAGNAGATCTGTAAC 120
 Qy 121 AATATGTACCTTCTCAGAAAAAATACAGTTTGTAAAGTAGATAAACTGGTTCTTTC 180
 Db 121 AAGTACGTGCGGACTTACCAAAACGCACACAACTGAAATTTGATAAAATCCATCGTT 180
 Qy 181 ATTATTTCCCTTGTGGATCCCTGGCGTTTGTAGTCACTTTGGACGTAATAGCC 240
 Db 181 ATCATTTCTGTAAGGATGTTCCGATGCTTTATTTGAGTTAGACGAGAAATTAAT 240
 Qy 241 ATTCTTACATGGGATGGCTTGTAGTCTGACCTCAAGCATAGAA-----GCTATTGTT 294
 Db 241 CTCTTACATGGGATGGCTTGTAGTCTGCTCAAGTAGATAGAAAAATCGCTGATTT 300
 Qy 295 AATGTGCAACACATTTAAAAATTAACAGACTCAATGATGCAAGAGATCCCTTGGC 354
 Db 301 GATAATACTCTGAAAAAAGTGAATAATAGATTAAATACGCTAAGACGAGACCCCTCTTGA 360
 Qy 355 AATCTATGGACAGGTACATGGCTATTGACCGCTGCTCCCGTAGGACCGGTCACTGGC 414
 Db 361 AATCTGTTGGTGAACGATGAATATGGTTTCAGATCATACGACGAGCAACATAGCTATA 420
 Qy 415 ACTTTATATCTTTAGGGGCTGATAAAAAAGTAAAAATGCAACGAGAGCAACATAGCTATA 474
 Db 421 GGCAGTTTCTGAGTTTATCTAATAAGCAAGTGAAGNACACGCTGCTGAAGTTGTATA 480

QY	475	GCAATGGCGTCCGTTGGAGTAATGATTTGAAGAAATGTTATATATTGATTCGGGAAA	534	CC	cell-cell interactions in higher eukaryotes for the development of		
Db	481	TCCAATGGTGTGCTTGGAGTAAAGATTTAAAAAGTTTATTACATTTGATTCGCTGTT	540	CC	insecticides, therapeutics and pharmaceutical drugs. The invention		
QY	535	AGAAGCTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA	594	CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA		
Db	541	AGACAGTCTGATCAATTTGATTTGATGCAAAAATTTATCACITTTCTAACCGCAACCG	600	CC	sequences (ABL01840-ABL16175), and the encoded proteins		
QY	595	TTATTTACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGCTCAAAATGATGAC	654	CC	(ABB57737-ABB72072).		
Db	601	TTATTTACATTTGACAAACATGGAATTTATGGGTTCCACTGATGACAGACTATAGATGC	660	CC	The sequence data for this patent did not form part of the printed		
QY	655	GAGGTTAATTTATGGGTTGCCGTTTCCAGGACAGCGAATTTATTAATCAGTACCCAA	714	CC	specification, but was obtained in electronic format directly from WIPO		
Db	661	GAAGGAACTTATGGGTAGCCACATGTCAGGTGATAAAGTTTAAAAATTTGATACTAGT	720	CC	at ftp.wipo.int/pub/published_pct_sequences.		
QY	715	CAACCGGAAGTGTACTGGATACCGTAAATATACAGATCCTCAGGTCACCTCTGTAGCA	774	XX	Sequence 912 BP; 231 A; 246 C; 238 G; 197 T; 0 other;		
Db	721	ACTCCGGAACCTTACTTGGAAATGTCGAGATTCAGAGATTCAGAGATCAGTCACTGTC	780	Query Match	10.7%; Score 98.6; DB 23; Length 912;		
QY	775	TTTGGCGGTCGGAATTTGGATGAATGATGATGATGATGATGATGATGATGATGATGAT	834	Best Local Similarity	49.2%; Pred. No. 1.7e-18;		
Db	781	ATTGGTGGAGCGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG	838	Matches	384; Conservative 0; Mismatches 379; Indels 18; Gaps 4;		
QY	835	AGTTCTTTNGACAAAGCTTTAGTTAATGGCAGCTCTACAGATAACAGGTTTAGGGCTC	894	QY	47	TTTGACAAGGTCTCACTGGGATCATGAACCTTATATATTTTCGTCGACACCGTAG	106
Db	839	-GTGCTGATGAACAAACCAATGAAGGGTCTATCTATAAGTAACCTGGACTTGGTG	897	Db	44	.TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTAGTGGACCTCGAAT	103
QY	895	AAAGTTTTCGCGGAGTTAAAGTGAAGCT	923	QY	107	AGAAAACTTTTCATAATATATGTACCTTCTCAGAAAAAATACACGTTTCTGTA	160
Db	898	AAAGGATTCGCGAGGATCGAGTTAAGTT	926	Db	104	CCCGCGCATTAATCGTTATGATTTCAAGCAGAAACAAAGTGTACAGGGCTAAAAATCGAGG	163
RESULT 4				QY	161	TAGATAAACTGGTTTCTTTCATTTATTTCCCTTCTGCTGGATCCCTGCGCGTTTGTAGTCA	220
ABL12833				Db	164	GCAGATATTTGCATCGTTTCATTTCTGCGGTTTGAGAACAAACCGCAGGAGTTTGC	223
ID	ABL12833	standard; cDNA; 912 BP.		QY	221	GTTTGGACGTGAATAGCCATTTTACATGGGATGCGGTTACTGTGCACTTACAAGCA	280
XX	AC	ABL12833;		Db	224	GATCGGCTCTTCGTACGGTCTCGTCAAGTGGGATGAGTCTCCGAGTGGCCNAGTCA	283
XX	DT	26-MAR-2002 (first entry)		QY	281	TAGAAGCTATTGTTAATGTCGAACACACATTTAAAAATACAGACTCAATGATGGCAAG	340
XX	DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 32981.		Db	284	CTCGACCCCTGTTTCGAGGTGCAGCGACCTGAAGGAAAAACCGCTTAATGATGCCAAA	343
XX	KW	Drosophila melanogaster expressed polynucleotide; insecticide;		QY	341	CAGATCCCTTGGCAATCTATGACAGGTGATGCTATTGACGCTGGTCTCCCGTAG	400
XX	KW	pharmaceutical; gene; ss.		Db	344	CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATTTCA	400
XX	OS	Drosophila melanogaster.		QY	401	GACCGGTCACTGGCAGTTTATATCATTTAGGGCTGATAAAAAAGTAAAAATGCAGAGA	460
XX	PN	WO200171042-A2.		Db	401	CCCAATGGAAGGTGAGCTCTACAGCTGGCAGCGCGGTGGACAGCCCAACGCTATCCGT	460
XX	PD	27-SEP-2001.		QY	461	GCAACATAGCTATAGCAATGGGCTCCGCTGGAGTAAATGATTTTCAAGAAATGATTTATA	520
XX	PF	23-MAR-2001; 2001WO-US09231.		Db	461	GCAAGTGGGCATATCCAATGGCTGGCTGGGATGTCNAAGGCCAAGAAAGTCTTACTTCA	520
XX	PR	23-MAR-2000; 2000US-191637P.		QY	521	TTGATTCCGGGAAAAGAGAGTAGACGAGTACGATTATGATGCTTCTACATTTATCCATCA	580
XX	PR	11-JUL-2000; 2000US-0614150.		Db	521	TCGACACCAACACACACGAGGTATTGGCCTATGACTACANTCAGAGCACCAGCCCGCTAA	580
XX	PA	(PEKE) PE CORP NY.		QY	581	GCAATCAACGGCCATTATTTACTTTTGAAAAA-----GCATGAAGTCCCTGGATATCCAG	634
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;		Db	581	GCAACCCNAAGGTCTATCTTCGATCTGAGGAAGATTTCGGCCGAGAGGACCATTTGTTCC	640
XX	PI	WPI; 2001-656860/75;		QY	635	ATGTCACCAATTTGATGAGAGGTAAATTTATGGGTTGCGGTTTCCAGGACAGCGAA	694
XX	PI	P-PSDB; ABB68730.		Db	641	ATGGCATGACCGTAGACACACCGATGCAATATCTACGTGGCCACCTTCAATGGTGGCAC	700
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more		QY	695	TTATTAATCAGTACCCCAACACCGGAGTGTACTGGATACCGTAAAAATACAGATC	754
XX	PT	genes from Drosophila and for elucidating cell signalling and cell-cell		Db	701	TCTTCAAGGTCAATCCAGACCGGTAAATTTCTGCTGGAGA---TCANAATTCNAACCA	757
XX	PT	interactions -		QY	755	CTCAGGTCACTCTGTAGCATTTGGCGCTCCGAATTTGGATGAGTCACTGATACATCTG	814
XX	PS	Claim 1; SEQ ID NO 32981; 2lpp + Sequence Listing; English.		Db	758	CCCAATCACCTCGGTGGCTTTTGGAGTCCCAATTTGGATATTTTGTATGTGCAACCG	817
XX	CC	The invention relates to an isolated nucleic acid detection reagent		QY	815	C 815	
XX	CC	capable of detecting 1000 or more genes from Drosophila. The invention is		Db	818	C 818	
XX	CC	useful in developmental biology and in elucidating cell signalling and					RESULT 5

Fri Aug 1 16:48:47 2003

us-10-089-986-1.rng

Query Match 8.7%; Score 80.8; DB 24; Length 1573;
Best Local Similarity 49.3%; Pred. No. 3.6e-13;
Matches 301; Conservative 0; Mismatches 298; Indels 12; Gaps 3;

QY 298 GTCGAACACACATTAATAATACAGACTCAATGATGGCAACAGACATCCCTTGGCAAT 357
DB 349 GTGGATGAAGATAAGAAAATATCGATTCAATGATGGGAAGGTGGATCTCGTGGAGA 408

QY 358 CTATGGACAGGTACAAATGATGACGCTGGTCTCCCGTA---GGACCGGTCACTGGC 414
DB 409 TACTTTGCTGATACATGGCTGAGGAACGCCGCCAGCTGTCTTGAGCGGCACCAAGG 468

QY 415 AGTTTATATCATTTAGGGGCTGATAAAAGTAAATAATGCAGAGACACATAGCTATA 474
DB 469 TCCCTGTACTCCCTCTTCTGATCAGATGTGAAGAAATACTTTGACCAAGTGGATATC 528

QY 475 GCAATGCGCTCGCTGAGTAATGATTGAAGAAATGATATATATTGATTTCGGGAAA 534
DB 529 TCAATGCTTTGGATTGCTCCCTGGACCATAAATCTTCTACTACATTGACAGCCTGTC 588

QY 535 AGAAGAGTAGAGAGTACGATTATGATGCTTCTTACATTATCCATCAGCAATCAACGCCA 594
DB 589 TACACTGTGGATGCTTTTGAATATGACCTACAAACAGGACAGATTTCCACCCGCAATT 648

QY 595 TTATTTACTTTGAAAGCATGAGTGCCTGGATATCCAGATGTCAACAAATGATGAG 654
DB 649 GTTTACAAGATGGAAGAAAGATGAA-----CAATCCAGATGGAATGCAATTGATGCT 702

QY 655 GAGGTAATTTATGGTGGCTTTTCCAGGACAGCGAATTTATAAATCAGTACCCAA 714
DB 703 GAGGAAAGCTATGGTGGCTGTACAAATGAGGAAGAGTAATTCGCTGGATCCTGAG 762

QY 715 CAACCGGAAGTGTACTGGATACCGTAAATATACAGATCCTCAGGTCACTCTGTAGCA 774
DB 763 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCCTGTGTATATAAACAATCTCATGCTG 819

QY 775 TTGGCGGTCCGAAATTTGGATGAATGATGATTAACATCTGCTGTCTTACGCTTTGAGC 834
DB 820 TTGGAGGAAAGATTACTCTGAATGATGTGACCTGTGCCAGGGATGGGTGAATGCT 879

QY 835 AGTTCTTTNGACAAGAATTTAGTTAATGGGACGCTCTACAGAGTAACAGGTTTGGCGTC 894
DB 880 GAAGGCGCTTTTGAGGCGAGCTGATGCTGGTAACATTTTCAAGATAACAGGCTCGGAGTC 939

QY 895 AAAGGTTTCGC 905
DB 940 AAAGGAATTGC 950

RESULT 7
AAQ90035
ID AAQ90035 standard; cdna to mRNA; 1356 BP.

XX AC AAQ90035;
XX DT 03-JAN-1996 (first entry)
XX DE Human SMP30 gene.
XX KW SMP30; marker protein; ageing; organ development; ss.
XX OS Homo sapiens.

XX Key Location/Qualifiers
XX FH 94..993
XX CDS /*tag= a
XX ET /product= SMP30 marker protein

XX JP07097399-A.
XX PN 11-APR-1995.
XX PD 29-SEP-1993; 93JP-0265681.
XX PF

XX 29-SEP-1993; 93JP-0265681.
XX (FJRE) FUJI REBIO KK.
XX WPI: 1995-175363/23.
XX P-PSDB; AAR74219.
XX Novel polypeptide for detecting human ageing marker protein SMP30 -
XX for monitoring liver and kidney development in new-born babies
XX ClalM 3; Page 6-7; 10pp: Japanese.
XX AAQ90035 is the SMP30 gene. It encodes the human ageing marker protein,
XX SMP30 (AAR74219). Human SMP30 is found in human organs, tissues, blood,
XX urine and cerebrospinal fluid. The blood concentration of SMP30 is
XX known to increase with renal and hepatic deficiencies and to decrease
XX with age. It is therefore useful in the monitoring of renal or hepatic
XX deficiencies and for the monitoring of the development of the liver
XX and kidneys in newborn babies.
XX Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;
SQ

Query Match 8.1%; Score 74.4; DB 16; Length 1356;
Best Local Similarity 48.6%; Pred. No. 2.5e-11;
Matches 297; Conservative 0; Mismatches 302; Indels 12; Gaps 3;

QY 298 GTCGAACACACATTAATAATACAGACTCAATGATGGCAACAGACATCCCTTGGCAAT 357
DB 370 GTGGATAACGACANGAANAACATCCCTTCAATGATGGGAAGTGGATCCCGCGAGG 429

QY 358 CTATGACAGGTACAATGGCTATTGACGCTGGTCTCCCG---TAGGACCGGTCACTGCG 414
DB 430 TACTTTGCTGGCACCATGGCTGAGGAACAGCTCCAGCAGTCTTTGAGCGGCACAGGG 489

QY 415 AGTTTATATCATTTAGGGCTGATATAAAGTAAAAATGCAGCAGCAACATAGCTATA 474
DB 490 GCCTGTACTCCCTCTTCTCTGATCACCAGCTGAAAAAGTACTTTTACCAGGTGGACAT 549

QY 475 GCAATGGGCTCGGTGGAGTAATGATTTGAAGAAATGATATATATATGATTTCGGGANA 534
DB 550 TCCAATGGTTGGATTGGTGGCTAGACCACAAATCTTCTATTACATGACAGCCTGTCC 609

QY 535 AGAAGAGTACAGAGTACGATTATGATGCTTCTACATTTCCATCAGCAATCAACGCCA 594
DB 610 TACTCCGTGGATGCTTTGACTATGACCTGCACAGGACAGATCTCCACCCGCAAGT 669

QY 595 TTATTTACTTTTGAAGAGTGAAGTGCCTGGATATCCAGATGGTCAACAAATGATGAG 654
DB 670 GTTTACAAGCTAGAAAAGGAAGAA-----CAATCCAGATGGAATGTATTGATGCT 723

QY 655 GAGGTAATTTATGGTTGGCTTTCCTTCCAGGACAGCGAATTTATAATCAGTACCCAA 714
DB 724 GAGGGAAGCTCTGGGTGGCTGTACATGGAAGAGAGTATCGTTTAGATCTCTGTG 783

QY 715 CAACCGGAAGTGTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACTCTGTAGCA 774
DB 784 ACAGGGAAA---AGACTTCAAACTGTGAAGTTGCCTGTGATAAACAACACTTCATGCTG 840

QY 775 TTTGGCGTCCGAATTTGGATGAACTGCATGTAACATCTGCTCTTCACTTACGTGACGAC 834
DB 841 TTTGGAGGGAAGAAATTAATCTCTGAAATGATGATGACCTGCGCGCGGATGGATGGACCCC 900

QY 835 AGTTCTTTNGACAAAAGTTTAAATAGGCGACCTCTACAGAGTAACAGGTTTAGGCGTC 894
DB 901 GAGGTCCTTTGAGGCAACCTGAAGCTGGTGGTAAATTTTCAAGATAACTGCTCGGGGTC 960

QY 895 AAAGCTTTCGC 905
DB 961 AAAGGAATTGC 971

RESULT 8

AAX40196
 ID AAX40196 standard; DNA; 1356 BP.
 XX AC AAX40196;
 XX DT 02-JUL-1999 (first entry)
 XX DE Senescence marker protein SMP30 gene.
 XX KW Cancer associated antigen; diagnosis; research; treatment; human;
 XX KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 XX KW prostate cancer; ss.
 XX OS Homo sapiens.
 XX PN W09904265-A2.
 XX PD 28-JAN-1999.
 XX PF 15-JUL-1998; 98WO-US14679.
 XX PR 22-JUN-1998; 98US-0102322.
 XX PR 17-JUL-1997; 97US-0896164.
 XX PR 10-OCT-1997; 97US-0061599.
 XX PR 10-OCT-1997; 97US-0061765.
 XX PR 10-OCT-1997; 97US-0948705.
 XX PR 11-OCT-1997; 97GB-0021697.
 XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 XX PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
 XX PI Tureci O;
 XX DR WPI: 1999-132448/11.
 XX P-PSDB; RAY06995.
 XX PT New isolated cancer associated nucleic acids and polypeptides -
 XX PT isolated using sera from cancer patients, used to develop products
 XX PT for the diagnosis, monitoring or treatment of cancers
 XX PS Claim 67; Page 775; 787pp; English.
 XX CC The invention relates to a method for diagnosing a disorder characterised
 XX CC by expression of a human cancer associated antigen precursor coded for by
 XX CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 XX CC biological sample isolated from a subject with an agent that specifically
 XX CC binds to the NAM, an expression product or a fragment of an expression
 XX CC product complexed with an HLA molecule; and (b) determining the
 XX CC interaction between the agent and the NAM or the expression product as a
 XX CC determination of the disorder. The products and methods can be used in
 XX CC the diagnosis, monitoring, research, or treatment of conditions
 XX CC characterised by the expression of various cancer associated antigens.
 XX CC The invention provides nucleic acid sequences and encoded polypeptides
 XX CC which are cancer associated antigen precursors expressed in human breast
 XX CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 XX CC lung cancer.
 XX SQ Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;
 Query Match 8.1%; Score 74.4; DB 20; Length 1356;
 Best Local Similarity 48.6%; Pred. No. 2.5e-11;
 Matches 297; Conservative 0; Mismatches 302; Indels 12; Gaps 3;
 QY 298 GTCCGACACACATTAAATAACAGACTCAATGTCGAAGCAGATCCCTTGGCAAT 357
 DB 370 GTGGATACGACAGAAAACAATCGCTCAATGATGGAAGGTGATCCCGCGGGAGG 429
 QY 358 CTATCGACAGGTACAATGGCTATTGACGCTGCTCTCCCG--TAGGACCGGTCACTGGC 414
 DB 430 TACTTTGTCGCCACCATGGCTGAGGAACACCTCCAGCAGTCTTGACGGCACCAGGGG 489
 QY 415 AGTTTATATATTTAGGGGCTGATAAAAGGTAATAATGCACGAGCAACATAGCTATA 474

Db 490 GCCCTGTACTCCCTCTTTCCCTGTATCACCACCTGAAAAAGTACTTTGACCGAGGTGGACATT 549
 QY 475 GCAAAATGGCTCGCTGAGTAATGATTGTAAGAAAAATGATTATATATTGATTCGGGAAA 534
 Db 550 TCCAAATGGTTGGATTGGCTCGCTAGACCACAAAATCTTCTATTACATTGACAGCTGTCC 609
 QY 535 AGAAGATGACAGTAGAGTATGATGCTTCTTACATTATCCATCCATCAGCAATCAACGGCCA 594
 Db 610 TACTCCGTGGATGCTTTGACTATGACCTGCAGACAGGACAGATCTCCAACCCGAGAAGT 669
 QY 595 TTATTTACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGGTCAAAACAATTCATGAG 654
 Db 670 GTTTACAAGCTAGAAAAAGGAAGAA-----CAAAATCCACAGATGGAATGTGATTGATGCT 723
 QY 655 GAGGTAAATTTATGGTTGGCTTTTCCAAAGCAGACAGCAATTTATTAATCAGTACCCAA 714
 Db 724 GAGGGAAAGCTCTGGTGGCTGTACAAATGGGAAGAGTGTCTGTTTATAGATCCTGTG 783
 QY 715 CAACCGGAAGTGTACTGTGATACCGTAAATAACAGATCCTCAGGTCACTCTGTAGCA 774
 Db 784 ACAGGAAA---AGACTTCAAACTGTGAAGTTGCTGTGATATAAAACAACCTTCATGTGTC 840
 QY 775 TTTGGCGTCCGAATTTGGATGAACATGATGATGATGATGATGATGATGATGATGATGATG 834
 Db 841 TTTGGAGGAAGAATTTACTTGAATATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 835 AGTTCTTTNGACAAAAGTTTAAATGAGGACGCTCTACAGATTAACAGGTTTAGGCGTC 894
 Db 901 GAGGTCTTTTGGAGCAACCTGAAGCTGGTGAATTTTCAAGATAAATCTGGTCTGGGGTC 960
 QY 895 AAAGTTTTCG 905
 Db 961 AAAGGAATTGC 971

RESULT 9
 ABN95150
 ID ABN95150 standard; DNA; 1356 BP.
 XX AC ABN95150;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #1648 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
 XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX PN W0200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX PI WPI: 2002-426119/45.
 XX DR Diagnosing and detecting the progression of liver cancer,
 XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 XX PT involves detecting the level of expression of two or more genes in a
 XX PT liver tissue sample
 XX PS Claim 1; SEQ ID NO 1648; 298pp; English.

CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;

Query Match 8.1%; Score 74.4; DB 24; Length 1356;
Best Local Similarity 48.6%; Pred. No. 2.5e-11;
Matches 297; Conservative 0; Mismatches 302; Indels 12; Gaps 3;

Qy 298 GTCGAACACACATTAATAAATACAGACTCAATGATGGCAAGCAGATCCCTTGGCAAT 357
Dy 370 GTGGATACGACAGAAACAAATCGCTTCAATGATGGAAGTGGATCCCGCGGGAGG 429
Qy 358 CTATGGAGACATGATGGCTTATGACGCTGGTCTCCCG---TAGACCGGTCACTGGC 414
Dy 430 TACTTTGTGGCACCATGGCTGAGGAACACAGCTCCAGCAGTCTTGAGCGGCACAGGGG 489
Qy 415 AGTTTATATCATTTAGGGGCTGATAAAGGTAAATGTCACGAGACACATAGCTATA 474
Dy 490 GGCCTGTACTCCTCTTCCCTGATCACCAGCTGAAAGTACTTTGACCGGTGACATT 549
Qy 475 GCAATGGGCTCGCTGGAGTAATGATTTGAAGAAATGATTAATATGATTCGGGAAA 534
Dy 550 TCCATAGTTGGATGTGCTGCTAGACCAACAAATCTTCTATTACATTGACAGCCTGTC 609
Qy 535 AGAAGAGTAGAGGAGTACGATATGATGCTTCTACATTCATCCAGCAATCAACGGCA 594
Dy 610 TACTCCGGTGGATGCTTTGACTATGACCTGCAGACAGACAGATCTCCACCGCAGAGT 669
Qy 595 TTATTACTTTTGAAGACGATGAGTGCCTGGATATCCAGATGTCACCAATGATGAG 654
Dy 670 GTTTACAGCTAGAAAGGAGAA-----CAAAATCCAGATGGAATGTGTATGATGCT 723
Qy 655 GAGGTAATTTATGGTGGCTTTTCCAGACAGAGCAATTAATAATCAGTACCA 714
Dy 724 GAGGGAAGCTCTGGGTGGCCTGTTTACAATGGAGGAGAGTGTGTTTATAGCTCTGTG 783
Qy 715 CAACGGGAAGTGTACTGGATACCGTAAATAATACAGATCCCTCAGGTCACTCTGTAGCA 774
Dy 784 ACAGGGAAA---AGACTTCAACTGTGAAGTGGCTGTGTATAAACAATCTCATGCTGC 840
Qy 775 TTGGCGGTCCGAATTTGGATGAATGATGATTAACATCTGGTGGTTCACGTTGACAC 834
Dy 841 TTGGAGGGAAGAAATTAATCTCTGAAATGATGTGACCTGGCCCGGGATGGATGACCCC 900
Qy 835 AGTCTTTNGACAAAGATTTAGTTAATGGCAGCTCTACAGAGTAAACAGTTTGGCGCTC 894
Dy 901 GAGGGTCTTTTGGGCAACCTGAAGCTGTGGTGAATTTTCAAGATAACTTGGTCTGGGGTC 960
Qy 895 AAAGTTTTCGC 905
Dy 961 AAAGGAATTC 971

RESULT 10
ID ABL12832/c
ID ABL12832 standard; cDNA; 2970 BP.
XX
AC ABL12832;
XX

DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 32978.
DE
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX P-PSDB; ABB68729.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions
XX
XX Claim 1; SEQ ID NO 32978; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2970 BP; 893 A; 558 C; 574 G; 945 T; 0 other;

Query Match 7.6%; Score 70.4; DB 23; Length 2970;
Best Local Similarity 47.3%; Pred. No. 5.3e-10;
Matches 320; Conservative 0; Mismatches 341; Indels 15; Gaps 3;

Qy 47 TTGAGAGAGTCTCTCACTGGGATCATGAACCTCAGACCTTATATTTCTGCGACACCGTAG 106
Dy 1927 TGGCGGAGGACCCCACTGGGATGTTGATCGCAGAGTCTGTACTAGTGGACCTCGAAT 1868
Qy 107 AGAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACAGTTTGTGTA---G 160
Dy 1867 CCGCGCGCATTAATCGTTATGATTTCAAGCAGAAACAAAGTGTACAGGGCTAAATCCAGG 1808
Qy 161 TAGATAAATCGTTTCTTTCATTAATCCCTTGTGATCCCTGCGCGCTTTTGTAGTCA 220.
Dy 1807 GCGAGATATTGCAATCTGTTTCTGCGGTTGGAGAAACAAACCCAGAGAGTTTGGCGTAG 1748
Qy 221 GTTTGGAACGTGAATAGCCATCTTACATGGATGGCGTTAGTGTGCACTACAAAGCA 280
Dy 1747 GATCGGCTTCTGTTACGGTCAATCCATCGGATGGGATGGAGTCTCCGCAAGTCA 1688
Qy 281 TAGAAGCTATTGTTAATGTCGAACACACATTAATAAATAACAGACTCAATGATGGCAAG 340
Dy 1687 CTCGACCCCTGTTTCAGGTGCGAGCGGACCTGAAAGAAACCCGCTTTAATGATGCCAAA 1628
Qy 341 CAGATCCCTTGGCAATCTATGGACAGGTACATGGCTATTGAGCGCTGGTCTCCCGCTAG 400
Dy 1627 CCGATCCCAATGGCGGTTTACGGTGGCACCACCTGGC---CGACAGTGGCGCACATATCA 1571

QY 401 GACCGGTCTACTGGCAGTTTATATCATTTAGGGCTGATATAAAGGTAATAATCCACGAGA 460
Db 1902 TCCGTCACT---GGCGAGCTGTACCGCTGGGAGCGGTGGCCAGGTGTCGGTGTCAAGG 1958
QY 451 CCAACATAGCTATAGCAANTGGCTCCGCTGGAGTATGATTTGAAGAAATGTTATTATA 520
Db 1959 CGATGTGGGCATCTCCAAATGGACTGGCATGGGACGAGAGGCAAGATTCCTACTACA 2018
QY 521 TTGATTCGGGAAAAGAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db 2019 TCGATACACCGACTACGAGTGAAGTCTGATGACTATGATTTCCGAGACGGCGTGCTA 2078
QY 581 GCANTCAACGGCCATTTACTTTTGAAGCAT-----GAAGTCCCTGGATATCCAG 634
Db 2079 GCAATCCCAAGGTTATATTCAACTGCGCAAGATAGTCCCAAGGATCATCTGCTGCCCG 2138
QY 635 ATGGTCAACAAATTTGATGAGGAGGTAAATTTATGGTTGGCGTTTTCCAAAGG 686
Db 2139 ATGGCTGACCATCGATACCGAGGCAACCTGTATGTGGCCACCTTCAATGG 2190

RESULT 15

ABL03466
ID ABL03466 standard; cDNA; 4624 BP.

XX ABL03466;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 4880.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB59363.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Claim 1; SEQ ID NO 4880; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 4624 BP; 1242 A; 1013 C; 1017 G; 1352 T; 0 other;

Query Match 5.7%; Score 52.8; DB 23; Length 4624;
Best Local Similarity 46.0%; Pred. No. 9,3e-05;
Matches 300; Conservative 0; Mismatches 337; Indels 15; Gaps 3;
QY 47 TTGAGAAAGGTCTCTCCTGAGGATCATGAACTCAGACCTTATATTTCTGCACACCGTAG 106
Db 2612 TGGCGAGGGTCCCATTTGGGATGTGGCGACGAGCGCTGTACTACGTGGATTTGGAGG 2671
QY 107 AGAAACATTTTCATATAATGATGCTCTCAGAAAA-----AATACACGTTTCTTAAG 160
Db 2672 CAGGACGCGCTGCTCCGCTACGACTATGCGCAAGAACAGGTCTACAAGACAAGATCGAGG 2731
QY 161 TAGATAAACTGCTTTCTTCATTATTCCTCTTGCTGGATCCCTCGCCCTGGCCGTTTGTAGTCA 220
Db 2732 GCGAAACCTTTGCGCGGATTCGTGCTGCGGTGGAGGAGCTCCGAGGAATTTGCGCGTGG 2791
QY 221 GTTTGGAAGCTGAATAGCCATTCTTACATGGGATGGCGTTAGTCTGCACCTACAAGCA 280
Db 2792 GCTGCGGTGCGACGCGTGGTGTGCTCACTGGGATGGGCTCTCGCCACGCGCAAGGTGG 2851
QY 281 TAGAAGCTATTGTTAATGTCGAACACACATTTAAAAATTAACAGACTCAATGATGGCAAG 340
Db 2852 TGGCGACACTGTTTCGAGGTGCGCCACTGATGGAGAAGAAATCGTTTGAACGACGCAAGG 2911
QY 341 CAGATCCCTTGGCAATCTATGGACAGGTACAAATGGCTATTGACGGTGGTCTCCCGCTAG 400
Db 2912 TTTGATCCCGTGGTCTGCTTTTGGCGGCACCATGCGCTTACATTTGGCGATGAGTTGAGT 2971
QY 401 GACCGGTCACTGGCAGTTTATATCATTTAGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
Db 2972 TCCGTCACT---GGCGAGCTGTACCGTGGGAGCGCGTGGCCAGGTGTCGGTGTATCAAG 3028
QY 461 GCAACATAGCTATAGCAAAATGGCTCGCGTGGAGTAAATGATTTGAAGAAAAATGTTATATA 520
Db 3029 GCGATGTGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAAGTTCTACTACA 3088
QY 521 TTGATTCGGGGAAGAAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db 3089 TCGATACCAACCGACTACGAGGTGAAGTGTGATGCTATGATTTTCGAGACCGGGGTGGCTA 3148
QY 581 GCAATCAACGGCCATTATTACTTTTGAAGAGCAT-----GAAGTCCCTGGATATCCAG 634
Db 3149 GCAATCCCAAGTTATATTCAATCTGCGCAAGAAATAGTCCCAAGGATCATCTGCTCCCG 3208
QY 635 ATGGTCAACAAATTTGATGAGGAGGTAATTTATGGGTTGCCGTTTTCCAAAGG 686
Db 3209 ATGGCTGACCATCGATACCGAGGCAACCTGTATGTGGCCACCTTCAATGG 3260

Search completed: August 1, 2003, 12:06:43
Job time : 305 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:01:54 ; Search time 2636 Seconds
(without alignments)
10201.439 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagtggttgtaaaa.....cgggagttaaagtaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_ov:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htg_hum:
40: em_htg_mus:
41: em_htg_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	923	99.9	927	3	AB062786	AB062786 Photinus
2	325.2	35.2	924	3	AB072447	AB072447 Luciola 1
3	320.4	34.7	930	3	AB072448	AB072448 Luciola c
4	98.6	10.7	1040	3	AY071130	AY071130 Drosophil
5	97.2	10.5	1150	3	AY028616	AY028616 Calliphor
6	93.8	10.2	1022	3	AB029490	AB029490 Drosophil
7	91	9.8	992	3	AF326959	AF326959 Calliphor
8	90.4	9.8	1216	6	E09237	E09237 cDNA insert
9	89.8	9.7	1155	3	AB036903	AB036903 Sarcophag
10	88.8	9.6	1594	10	RAYCBP1	D38467 Rat mRNA fo
11	85.6	9.3	1594	10	RNSMP30A	X69021 R.norvegicu
12	85.6	9.3	1605	10	AB037934	AB037934 Rattus no
13	80.8	8.7	1109	10	D86217	D86217 Mus musculu
14	80.8	8.7	1573	10	MMU28937	U28937 Mus musculu
15	80.8	8.7	1598	10	BC012710	BC012710 Mus muscu
16	77.4	8.4	1126	5	AB037936	AB037936 xenopus 1
17	74.4	8.1	1356	6	AX409001	AX409001 Sequence
18	74.4	8.1	1356	6	E09013	E09013 cDNA encodi
19	74.4	8.1	1356	9	HUMSMP30	D31815 Human mRNa
20	74.4	8.1	1375	9	AB028125	AB028125 Homo sapi
21	74.4	8.1	1438	9	AB033064	AB033064 Homo sapi
22	73.6	8.0	1088	5	AB033368	AB033368 xenopus 1
23	73.2	7.9	269050	1	AP000987	AP000987 Sulfolobu
24	71.6	7.7	1394	4	AB035446	AB035446 Bos tauru
25	70.4	7.6	49817	2	AC013939	AC013939 Drosophil
26	70.4	7.6	167195	3	AC007808	AC007808 Drosophil
27	70.4	7.6	254961	3	AE003706	AE003706 Drosophil
28	69.4	7.5	888	1	SAU05664	U05664 Sulfolobus
29	67.8	7.3	1357	4	AB035445	AB035445 Oryctolag
30	67	7.3	1108	3	AY118643	AY118643 Drosophil
31	67	7.3	1208	3	AB036904	AB036904 Drosophil
32	65	7.0	56152	2	AC116963	AC116963 Dictyoste
33	62.8	6.8	633	3	AB079885	AB079885 Sarcophag
34	62.8	6.8	10105	1	AE008864	AE008864 Sulfolobu
35	59.8	6.5	286756	1	AP000990	AP000990 Sulfolobu
36	59.2	6.4	1221	5	AB037935	AB037935 Gallus ga
37	52.8	5.7	76808	2	AC014187	AC014187 Drosophil
38	52.8	5.7	186714	3	AC023688	AC023688 Drosophil
39	52.8	5.7	301051	3	AE003487	AE003487 Drosophil
40	50.4	5.5	12835	1	AE008895	AE008895 Sulfolobu
41	46.6	5.0	24650	2	AC023731	AC023731 Drosophil
42	45.6	4.9	1141	6	AX083744	AX083744 Sequence
43	44.4	4.8	100886	2	AP003831	AP003831 Oryza sat
44	44.4	4.8	164237	2	AP005199	AP005199 Oryza sat
45	43.4	4.7	7218	6	I66494	I66494 Sequence 14

ALIGNMENTS

RESULT 1
AB062786
LOCUS
DEFINITION Photinus pyralis mRNA for luciferin regenerating enzyme.. complete cds.
ACCESSION AB062786
VERSION AB062786.1 GI:14331151
KEYWORDS Photinus pyralis
SOURCE Photinus pyralis cDNA to mRNA.
ORGANISM Photinus pyralis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia;
Cantharoidea; Lampyridae; Photinus.
REFERENCE 1
AUTHORS Gomi,K. and Kajiyama,N.

TITLE Oxyluciferin, a luminescence product of firefly luciferase, is enzymatically regenerated into luciferin
J. Biol. Chem. 276 (39), 36508-36513 (2001)
21453302
REFERENCE 2 (bases 1 to 927)
AUTHORS Gomi, K. and Kajiyama, N.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2001) Keiko Gomi, Kikkoman Corporation, Research & Development; Noda 399, Noda-shi, Chiba 2780037, Japan
(E-mail: 8553@mail.kikkoman.co.jp, Tel: 81471235579)

FEATURES
Location/Qualifiers
1..927
/organism="Photinus pyralis"
/db_xref="taxon:7054"
1..927
/codon_start=1
/product="luciferin regenerating enzyme"
/protein_id="BAB60700.1"
/db_xref="GI:14331152"
/translation="MGPPVEKIAELSKYTVGEGPHWDHETQTLFVDTVEKTFHKYVP
SOKRYTECKVDKLVSFIIPLAGSPGFVSLEREIALTLWDGVSAAPTSTIEALVNEP
HIKNRNLDKADPLNLTGMAIDAGLPVPTGLSLYHLGADKXKHSHESIAIAN
GLANNDLKKWYIDSKRRVDEYDASTLSISNORPLTFEKEHVPYDQGTIDE
EGNLWAVFOGRII KISTOOPEVLIDTVKIPDPOVTSVAFGGPNLDLHVTSAQLQL
DDSLDKSLNGHYRVYRTGLGVKGFAKVKL"
BASE COUNT 285 a 171 c 216 g 255 t

ORIGIN
Query Match 99.9%; Score 923; DB 3; Length 927;
Best Local Similarity 99.9%; Pred. No. 3e-220;
Matches 923; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGCCAGTGTGAAAAATTCGAGAACTTCGCAAGTATACGGTTGGAGAGGTCCT 60
DB 1 ATGGGGCCAGTGTGAAAAATTCGAGAACTTCGCAAGTATACGGTTGGAGAGGTCCT 60
QY 61 CACTGGGATCATGAACTCAGACCTTATATTCGTCGACACCGTAGAGAAAACCTTTTCAT 120
DB 61 CACTGGGATCATGAACTCAGACCTTATATTCGTCGACACCGTAGAGAAAACCTTTTCAT 120
QY 121 AAATATGATACCTTCACAGAAAATACACGTTTTGTAAAGTAGATAAATCTGTTCTTC 180
DB 121 AAATATGATACCTTCACAGAAAATACACGTTTTGTAAAGTAGATAAATCTGTTCTTC 180
QY 181 ATTATTCCTTCGCTGGATCCCTGGCGGTTTTGTAGTCAGTTTGGACGTTGAAATAGCC 240
DB 181 ATTATTCCTTCGCTGGATCCCTGGCGGTTTTGTAGTCAGTTTGGACGTTGAAATAGCC 240
QY 241 ATCTTACATGGATGGCGTTAGTCTGCACCTACAGCATAGAGCTATTGTTAATGTC 300
DB 241 ATCTTACATGGATGGCGTTAGTCTGCACCTACAGCATAGAGCTATTGTTAATGTC 300
QY 301 GAACACACATTAATAAATACAGCTCAATGATGCAAGCAGATCCCTTTGGCAATCTA 360
DB 301 GAACACACATTAATAAATACAGCTCAATGATGCAAGCAGATCCCTTTGGCAATCTA 360
QY 361 TGGACAGGTACAAATGGCTATTGACGCTGTTCTCCCGCTAGGACCGGTCACCTGGCAGTTA 420
DB 361 TGGACAGGTACAAATGGCTATTGACGCTGTTCTCCCGCTAGGACCGGTCACCTGGCAGTTA 420
QY 421 TATCATTTAGGGGCTGATAAAGGTAAATATGATGCTAGGACGAGCAACATAGCTATAGCAAT 480
DB 421 TATCATTTAGGGGCTGATAAAGGTAAATATGATGCTAGGACGAGCAACATAGCTATAGCAAT 480
QY 481 GGGCTCGCTGGAGTAATGATTGAAGAAATGTTATATTGATTCGGGGAAGAAGA 540
DB 481 GGGCTCGCTGGAGTAATGATTGAAGAAATGTTATATTGATTCGGGGAAGAAGA 540
QY 541 GTAGAGGATAGGATTTATGATGCTTCTACATTTATCCATCAGCAATCAACGGCCATTATT 600
DB 541 GTAGAGGATAGGATTTATGATGCTTCTACATTTATCCATCAGCAATCAACGGCCATTATT 600
QY 601 ACTTTTGAAGACGATGAAGTGGCTGGATATCCAGATGCTCAACAAATTTGATGAGGAGGT 660

DB 601 ACTTTTGAAGACGATGAAGTGGCTGGATATCCAGATGCTCAACAAATTTGATGAGGAGGT 660
QY 661 AATTTATGGTGGCGGTTTTCACAGGACAGGGAATTTAAATCAGTACCAACAACCG 720
DB 661 AATTTATGGTGGCGGTTTTCACAGGACAGGGAATTTAAATCAGTACCAACAACCG 720
QY 721 GAAGTGTACTGATACCGTAAATAATACAGATCCTCAGGTCACCTCTGTAGCATTTTGGC 780
DB 721 GAAGTGTACTGATACCGTAAATAATACAGATCCTCAGGTCACCTCTGTAGCATTTTGGC 780
QY 781 GGTCCGAATTTGGATGAATCAATCTAATCTGCTGGTCTTCAGCTTTCAGCAGCAATTTCT 840
DB 781 GGTCCGAATTTGGATGAATCAATCTAATCTGCTGGTCTTCAGCTTTCAGCAGCAATTTCT 840
QY 841 TTNGACAAAAGTTTAAATAGTGAAGCTA 924
DB 841 TTNGACAAAAGTTTAAATAGTGAAGCTA 924
QY 901 TTGCGGGAGTTAAAGTGAAGCTA 924
DB 901 TTGCGGGAGTTAAAGTGAAGCTA 924
RESULT 2
AB072447 924 bp mRNA linear INV 22-FEB-2002
LOCUS Luciola lateralis H-LRE mRNA for luciferin-regenerating enzyme,
complete cds.
ACCESSION AB072447
VERSION AB072447.1 GI:18857904
KEYWORDS Luciola lateralis cDNA to mRNA.
ORGANISM Luciola lateralis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia;
Cantharoidea; Lampyridae; Luciola.
REFERENCE 1
AUTHORS Gomi, K., Hirokawa, K. and Kajiyama, N.
TITLE Molecular cloning and expression of luciferin-regenerating enzyme (LRE) cDNA in Luciola cruciata and Luciola lateralis
JOURNAL Unpublished
2 (bases 1 to 924)
AUTHORS Gomi, K., Hirokawa, K. and Kajiyama, N.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Keiko Gomi, Kikkoman Corporation, Research and Development Division; Noda 399, Noda, Chiba 2780037, Japan
(E-mail: 8553@mail.kikkoman.co.jp, Tel: 81471235579, Fax: 81471235550)
FEATURES
Location/Qualifiers
1..924
/organism="Luciola lateralis"
/db_xref="taxon:7052"
1..924
/gene="H-LRE"
1..924
/gene="H-LRE"
/codon_start=1
/product="luciferin-regenerating enzyme"
/protein_id="BAB85478.1"
/db_xref="GI:18857905"
/translation="MSPTVEIQTEVDFQIGEGPHWDHETQTLFVDTVEKTFHKYVP
STQKHQKMLNKRPSFTIPKETSDRFVLSLRDICTVLTWDGVSATPSHLETIVTVD
GIGETFNDSKADAFGNLWAGTLYSKFDEKQGNPTGLYSLSNKOLRKHISNIFLSN
GLAWNDSKRFYIDSKNRITDQFDYDSENLIISNCQPLFTLDKHQIGLPLDQATIDE
NDNLWAVIRGKVINIGTKQPSLGLGVNPNESLITTSVCFGGSKLDELKLVYTSGLINE
YETDSTKLKVGGLYRVYRTGLGVKGLPAHREFSL"
BASE COUNT 325 a 146 c 178 g 275 t
ORIGIN

Query Match 35.2%; Score 325.2; DB 3; Length 924;
Best Local Similarity 60.2%; Pred. No. 1.1e-70;
Matches 556; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

Qy	1	ATGGGGCCAGTGTGTGAAAAAATTCAGAACTTGCAGAGTATACGGTTGGAGAAAGGTCT	60
Db	1	ATGTCGGCAGTTATTTGAACAGATCACTGAAGTAGACAATTTCCAAATCGGAGAGGTCCA	60
Qy	61	CAC TGGGATCATGAACCTCAGACCTTATATTTCTGTCGACACCGTAGAGAAACTTTTCAT	120
Db	61	CAC TGGGATACCGAACAACAAGTTGTGATTTTGGGATATTC TAGAAAAATCCATACAC	120
Qy	121	AAATATGTACCTTC TCAGAAAAATACACGTTTTTGTAAAGTAGATAAACTGGTTCTCTTC	180
Db	121	AAATACGTACCATCGACAAAAACAATACTAAATGATTTTGAATAAACGTCGCTTTT	180
Qy	181	ATTATATCCCTTGTCTGGATCCCTTGGCGGTTTTGTGTAGTCAGTTTGGAAAGTGAATAGCC	240
Db	181	ATTATACCAATAAAAGAAACATCTGTCGGTTTTGTCTAAGTTTAGAACGACACATTTGC	240
Qy	241	ATTCTTACATGGGATGGCGTTAGTGTCTCACCTACAAGCATAGAAGCTATTGTTAATGTC	300
Db	241	GTCTTACTTGGGATGGTTGTAGTGCACGCCAAGTCATTTAGAAACAATAGTTACCGTT	300
Qy	301	GAACCACACATTAATAATACAGACTCAATGATGCCAAAGCAGATCCCTTTGGCAATCTA	360
Db	301	GATACGGGAATCGAAGGAATACATTCATGATGTTAAGCAGATCGCTTTGGCAATTTG	360
Qy	361	TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCGCTAGACCGGTCACTGGCAGTTTA	420
Db	361	TGGCAGGTACATTTATAGTAAATTCGATATTGAAAAACAAGGTCTTAATACAGGAACA	420
Qy	421	TATCATTTAGGGCTGATAAAAGTTAAATATGCACGAGACAACATAGCTATACCAAT	480
Db	421	TTGTACAGCTGTCTAATAGCAGTTAGAAACAATATTTCTAATATCTTCTCTATCAAT	480
Qy	481	GGGCTCGCTGGAGTAATGATTTGAAGAAAAATGATATATTGATTCGGGGAAAAAGA	540
Db	481	GGCCTACCTGGATAAGACTCAAAAAATTTATTTATCGACTCCAATAAAGAACA	540
Qy	541	GTAGCAGGTACGATTAATGATCTTCTACATTATCCATCAGCAATCAACGGCCATTATT	600
Db	541	ATAGATCAGTTGATTTATGATCTTGAAAAATTTAATTTATCAAAATTTGTCAACCATTT	600
Qy	601	ACTTTTCAAAGCATGAAGTCCCTGGATATCCAGATGTCACAACAATTTGATGAGAGGGT	660
Db	601	ACTGTGACAAACATGGGATACAGGGTTTACTGTATGCCAAACGATAGATGAATAATGAT	660
Qy	661	AAATTATGGGTTGGCGTTTTTCCAAAGGACAGCGAATTTAAAAATCAGTACCCCAACCG	720
Db	661	AATTTATGGTGTCTATAGTTCGAGGAGAAAGTTATAATATTTGGTACTAAGCAACCC	720
Qy	721	GAAGTGTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACTCTGTAGCATTTGGC	780
Db	721	GAATCTTTGCTGTGTATTAAAGATGCCTGAAAGTTTTGATAAATTCAGTTTGTTTTGA	780
Qy	781	GGTCCGAATTTGGATGAACTCATGTACATCTGCTGCTCTACGTTTCAGCTTCACGACAGT	840
Db	781	GGATCAAGTTGGACGAACTTTATGTAACGACTTCTGGTATCAAGAAGTATGAAA--CT	837
Qy	841	TTTGACAAAAAGTTTATGTTAATGGGCACGCTCTACAGAGTAACAGGTTTAGCGCTCAAAGT	900
Db	838	GACTCTACAAAAC TAGTAAAGGGTGGATTGTACAGAGTTACTGGATTAGGTCTTAAGGT	897
Qy	901	TTCCGGGGAGTTAAAGTGAAGCT	923
Db	898	TTACCCGCACACAGATTCAAGTCT	920

RESULT 3	AB072448	930 bp	linear	INV 22-FEB-2002
LOCUS	AB072448			
DEFINITION	Luciola cruciata G-LRE mRNA for luciferin-regenerating enzyme, complete cds.			
ACCESSION	AB072448			
VERSION	AB072448.1			GI:18857906
KEYWORDS				

SOURCE	Luciola cruciata cDNA to mRNA.
ORGANISM	Luciola cruciata Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE	1
AUTHORS	Gomi,K., Hirokawa,K. and Kajiyama,N.
TITLE	Molecular cloning and expression of luciferin-regenerating enzyme (LRE) cDNA in Luciola cruciata and Luciola lateralis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 930)
AUTHORS	Gomi,K., Hirokawa,K. and Kajiyama,N.
TITLE	Direct Submission
JOURNAL	Submitted (04-OCT-2001) Keiko Gomi, Kikkoman Corporation, Research and Development Division; Noda 399, Noda, Chiba 2780037, Japan (E-mail:8553@mail.kikkoman.co.jp; Tel:81471295579, Fax:81471235550)
FEATURES	Location/Qualifiers
source	1. .930 /organism="Luciola cruciata" /db_xref="taxon:7051"
gene	1. .930 /gene="G-LRE"
CDS	1. .930 /gene="G-LRE" /codon_start=1 /product="luciferin-regenerating enzyme" /protein_id="BAB85479.1" /db_xref="GI:18857907" /translation="MAPTVBQIVELGTLLIAESPHMDEQSLYFVDIVGRSVNKYVPY TTTHTOLKDFKNPSFTIPYKGCSDRFTIVSLEREINLLTWDGASSAPSKTEKTAVFDD TPKSENRLNDGARDPLGNLWAGTMNGSDHHTCTPVRGTTLSLSNKQVHEIXVCIT SGLNWSKDLKKYYIDSAVRQDQDFDANKLSLSNRQLPFTFDKHKIGMSPDQGTI DAGNLNWAFTCCQGDGLKIDITSTPTELLGIVEIPEHQTSGIGGAELNLYVTTSI KLPCADMETPKMGAIKYVTGLGVKGLPGDRVKI"
BASE COUNT	305 a 150 c 203 g 272 t
ORIGIN	

Query Match	34.7%;	Score 320.4;	DB 3;	Length 930;
Best Local Similarity	60.6%;	Pred. No. 1.7e-69;		
Matches 563; Conservative	0;	Mismatches 357;	Indels	Gaps 2;
Qy	1	ATGGGGCCAGTGTGTGAAAAAATTCGACAACTTCGGCAAGTATACGGTTGGAGAAGGTCCT	60	
Db	1	ATGGCTCCAACTGTTGAACAATAAGTAGTGAATTCGGCACTTATTTGCTTCGAGAAAGTCCA	60	
Qy	61	CATCTGGGATCATGAACTCTAGACCTTATATTTCTCGACACCGTAGAGAAAATCTTTTCAT	120	
Db	61	CATTTGGGACGACGAAACTCAAAAGTTTCTACTTTGTGTGATATGTFAGAAAGATCTGTAAAC	120	
Qy	121	AAATATATGACTCTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTTC	180	
Db	121	AAGTACGTGCGGACTACCAAAACGCACACACAACTGAAATTTGATAAAAAATCCATCGTTT	180	
Qy	181	ATATATCCCTTGGCTGGATCCCCTGGCCGTTTTCGTAGTCAGTTTGGAAACGTGAATAGCC	240	
Db	181	ATCATCTCTGTCAAAGGATGTCGGATCGTTTATTTGTGATTTTGAACAGAGAAATTAAT	240	
Qy	241	ATTCCTTACATGGGATGGCGTTAGTGCCTGCACCTACAAGCATAGAA-----GCTATTGTGT	294	
Db	241	CTTCTTACATGGGATGGTGTAGTCTCTCCAGTAGATAGAAAAAATCGCTGTATTT	300	
Qy	295	AAATCGCAACACACATTAATAATAACAGACTCAATGATGCCAAGCAGATCCCTCTGGC	354	
Db	301	GATAAATACTCTCGAAAAAAGTGAATAATAGATTAAATGACGGTAAGCAGACGCCCTCTTGG	360	
Qy	355	AATCTATGGACAGGTACAAATGGCTATTTCAGCGTGGTCTCCCGTAGGACGGTCACCTGGC	414	
Db	361	AACTGTGGGCTGGAACGNTGAATATGGGTTTCAGATCATACGACAGAAACACCGCTTCGT	420	
Qy	415	AGTTTATATCATTTAGGGGCTGATAAAAAAGTTAAAAATGCAACGACGACCAATAGCTATA	474	
Db	421	GGACGTTGTGCACTTTATCTAATAAGCAAGTGAAGAACAACACTGTCTGAAGTTTGTATA	480	


```

Qy 635 ATGGTCAACAAATGATGAGGAGGTAATTTATGGGTTGCGGTTTCCAAAGGACAGCGAA 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 ATGGCATGAGCTAGACACCGGCAATATCTACGTGGCCACCTTCAATGGTGCCACCG 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 695 TTATTAATAATCAGTACCAACACCGGAAGTTTACTGGATACCGTAAATAATACAGATC 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 749 TCTTCAGGTCATCCAAAGCACCAGCGTAAATCTCTGGAGA---TCAAAATCCAAACCA 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 755 CTCAGGTACCTCTGTAGCAATTTGGCGGTCGGAATTTGGATGCACTGATCAATCTG 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 CCCAAATCACCTCGTGGCTTTTGGAGTCCCAATTTGGATATTTTGTATGTGCAACCG 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 815 C 815
Db 866 C 866

RESULT 5
AY028616
LOCUS AY028616 1150 bp mRNA linear INV 29-MAR-2001
DEFINITION Calliphora vicina anterior fat body protein mRNA, complete cds.
ACCESSION AY028616
VERSION AY028616.1 GI:13488604
KEYWORDS Calliphora vicina.
SOURCE Calliphora vicina.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Calliphoridae; Calliphora.
REFERENCE 1 (bases 1 to 1150)
Hansen, I.A., Meyer, S.R., Schaefer, I. and Scheller, K.
Identification of the anterior fat body protein as Interactor of
the hexamerin receptor in the blowfly Calliphora vicina
Unpublished
REFERENCE 2 (bases 1 to 1150)
Hansen, I.A., Meyer, S.R., Schaefer, I. and Scheller, K.
Direct Submission
AUTHORS Hansen, I.A., Meyer, S.R., Schaefer, I. and Scheller, K.
TITLE Submitted (17-MAR-2001) Cell and Developmental Biology, University
Wuerzburg, BioCenter Hubland, Wuerzburg D-97074, Germany
JOURNAL Location/Qualifiers
FEATURES
source
1..1150
/organism="Calliphora vicina"
/db_xref="taxon:7373"
/cell_type="trophocytes"
/tissue_type="anterior fat body"
/dev_stage="larvae"
42..962
/note="AFP; strong interactor with the hexamerin receptor
in a yeast-2-hybrid interaction screen; CVAFP"
/codon_start=1
/product="anterior fat body protein"
/protein_id="AAK26174.1"
/db_xref="GI:13488605"
/translacion="MSVKVPLPDSHAELGEGPHWDVATONLYVDINAGKLLRYNN
ENKVKAKISEGLAGFAVPVEGTTDFQFVSGRRVIVQDVGSEAKVLKTLFEVQ
TGDERFTGNFNDKCDPRGLFAGTIRVGDPEHRIGELYKIEVEVEIKSDVG
ISGLANWETKKFYIIDDYEVKDYDFETGKASNPVKVFNLRKTSKDHLLPDG
MTIDEGNLNKKFYIINPTGVQVLEIKFPCKQITSAAFGPNFHPILFVTS
SRFGEHPAGTTPKVTGLGAGVPMKIQI"
BASE COUNT 338 a 223 c 231 g 358 t
ORIGIN

Query Match 10.5%; Score 97.2; DB 3; Length 1150;
Best Local Similarity 50.9%; Pred. No. 1.2e-13;
Matches 399; Conservative 0; Mismatches 358; Indels 27; Gaps 6;

Qy 47 TTGAGAGGTCCTCACTGGGATCATGAACTCAGACCTTATATTTGTCGACACCGTAG 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 TAGCGAAGGACCTCATTTGGGATGTGCCACCCAGAAATTTGTACTATGTTGACATCAATG 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 107 --AGAAAATTTTCATAAATATGATGACCTTCTCAGAAAA---AATACACGTTTGTAAAG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 CCGCAAACTGTTGGTTACAACTACAAATGAAATAAAGTGTACAGGCCAAATTTGAAG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 161 TAGATAAATCGGTTCTTTTCATTTATCCCTTGTGATCCCTCGGCGTTTTCGTAGTCA 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 GTGAGGATTTAGCTGGTTTCTGCTGTCGCCGTGCAAGGTACAAACGATCAATTTCTGCTGC 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 221 GTTTGGACGTGAATAGCCATTTCTACATGGGATGGGTTAGTGTGTCACCTACAGCA 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 GTTCTGTGCTGCTGTTGTCATGTACAATGGGATGGTGTCTTCTGAGACCGCTAAAGTAT 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 281 TAGAAGCTTATTTAATGTGCAACCAACACATTTAAAAAT-----AACAGACTCAATG 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 TGAAGACTTTGTCGAAGTGCAAACCGGTGATGAACGTTTCACCTGGCAACCGTTTCAATG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 332 ATGCCAAAGCAGATCCCTTTGGCAATCTATGGACAGGTACAAATGGCTATTGACGCTGCTG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 ATGCCAAGTGTGATCCTCGTGGTGTGTTGTCGCCGTACCATCGCTTATATCTCGCGGATG 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 392 TCCCGGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGCTGATAAAAAGTAAAAA 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 AATTTCGAGCATCGTTA---TGGTGAATATATACAAGTACGAAATGGGAGCGGAAGTTGAGG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 452 TGCACGAGAGCAACATAGCTATAGCAAAATGGGCTCGGCTGGAGTAATGATTTGAAGAAAA 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 TAATCAATCGGATGTGGTATTTCCATGTTTGGCTTGGATGAGAAGACTAAGAAT 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 512 TGTATTATATGATTTCGGGGAAAAAGAGAGTAGACGAGTAGCATGATATGATGCTTACAT 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 TCTATTACATCGATACCAACCGATTATGAGGTAAAGGAGTATGATATGATTTTGAACCTG 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 572 TATCCATCAGCAATCAACGGCCATTTACTTTTGAAG-----CATGAAGTGCCTG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 GCAGGGCTTCAATCCCAAGTTGTATTCAATCTACGCAAGCAGACGCCCCAAGGATCAAT 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 626 GATATCCAGATGGTCAACAATTTGATGAGGAGGTTAATTTATGGTTGCCCTTTTCCAAAG 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 TGTTCGCCGATGGCATGACCAATGATCTAGGAGCAATCTTTATGTGGCCACTTTTCAATG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 686 GACAGCAATTTATTAATCAGTACCCCAACACCGGAAGTGTACTGTGATACCGTAAAAA 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 GCCATTCCATCTACAAAATTAATCCCACTGCGCAAGTTTGTGTGAAAA---TCAAT 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 746 TACCAGATCCTCAGTCACTCTGTAGCATTTTGGCGTCCCAATTTGGATGACATGCATG 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 TCCCTGCAAAACAAATCACCTCGGCTGCTTTGGTGGTCCCAATTTCCATATTTTGTGTTG 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 806 TAAC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 TAAC 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AB029490
LOCUS AB029490 1022 bp mRNA linear INV 03-MAY-2002
DEFINITION Drosophila melanogaster mRNA for senescence marker protein-30
(SMP-30), complete cds.
ACCESSION AB029490
VERSION AB029490.1 GI:7339661
KEYWORDS senescence marker protein-30 (SMP-30).
SOURCE Drosophila melanogaster (strain:Canton S) 7 day adult stage whole
body cDNA to mRNA.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
Goto, S.G.
Expression of Drosophila homologue of senescence marker protein-30
during cold acclimation
J. Insect Physiol. 46 (7), 1111-1120 (2000)
PUBMED 10817837
REFERENCE 2 (bases 1 to 1022)
Goto, S.G.
Direct Submission
AUTHORS
TITLE

```

Fri Aug 1 16:48:46 2003

JOURNAL Submitted (01-JUL-1999) Shin G Goto, Ohio State University,
Department of Entomology, 1735 Neil Avenue, Columbus, OH 43210, USA
(E-mail: shingoto@sci.osaka-cu.ac.jp, Tel.: +1-614-292-4477,
Fax: +1-614-292-2180)

FEATURES

Location/Qualifiers
1..1022
/organism="Drosophila melanogaster"
/strain="Canton S"
/db_xref="taxon:7227"
/tissue_type="whole body"
/dev_stage="7 day adult stage"
29..340
/codon_start=1
/product="senescence marker protein-30 (SMP-30)"
/protein_id="BAA92938.1"
/db_xref="GI:7339662"
/translation="MSVKVRAVPSYAALGEGPHWDVROSLYYVLESAGINRYDFK
ONKYKAKIEGEIFASFILPVENKPOEFVAVGCGRLTVIOMDGVSAVAKYRTRLFVQ
PDLKRNLDKATDPNGRFYGGTMAQSDIFITQWKGELYSHQAGGPNARISKVGISN
GLAWDVAKKFEYFDTNNHVEYADYNOSTGAVSNPKVIFDLRRKIRPEGLFFDGMIV
DTDGNIVATFNGTGVFKVNPSTGKILKEIKIPTQITTSVAFGPNLDILYVTTANKE
DQPKPAGTTFQVTLGNAGYAGVNLKI"
267 a 260 c 249 g 226 t

CDS

BASE COUNT 287 a 260 c 249 g 226 t
ORIGIN
Query Match 10.2%; Score 93.8; DB 3; Length 1022;
Best Local Similarity 48.8%; Pred. No. 8.3e-13;
Matches 381; Conservative 0; Mismatches 382; Indels 18; Gaps 4;

QY 47 TTGGAGAAGTCTCCTCAGTCGGATCATGAACCTACAGACCTTATATTTGCTGCACACCGTAG 106
DB 72 TCGGGGAGGCCACCTGGGATGTGTGATCGCCAGAGCTCTACTACGTGCAGCTCGAAT 131
QY 107 AGAAACATTTTCATAAATATGTACCTCTTCAGAAAATACACGTTTGTGTA-----G 160
DB 132 CGCGCGGATTAATCGTTACGATTTCAAGCAGACAGCAAGGTACAAAGCTAAAGCTAATCGAGG 191
QY 161 TAGATAAATCGTTTCTTTCATTATTCCTTTCCTGATCCCTGCGGCTTTTGTAGTCA 220
DB 192 GCGAGATATTCATCTCTTCATCTCGCGGTGTGAGAACAAACCGCAGGAGTTGCGGTAG 251
QY 221 GTTTGAAGCTGAATAGGATCTTACATGATGGATGGGCTTAGTCTCCACCTACACGA 280
DB 252 GATCGGCTCTCGTACGCTCAGTCCAGTCCAGTGGAGTCTCCGCACTGGCCCAAGGTCA 311
QY 281 TAGAAGCTATTGTTATGTCGACACACACATTAATAAATACAGACTCAATGATGCAAG 340
DB 312 CTCGCAACCTGTTGAGGTGCGACCGGATCTCGAAGAAACCGCCTTAATGATGCCA 371
QY 341 CAGATCCCTTGGCAATCTATGGACAGGTACAAATGGCTATGACGCTGTCTCCCGGTAG 400
DB 372 CGGATCCCAATGGCGCTTTTACGGTGGCACCACCATGGC---CGACAGTGGCGATATCTTCA 428
QY 401 GACCGTCTACTGCGAGTTTATATCATTTAGGGCTGATAAAGGTAAATAATGACACGAGA 460
DB 429 CCCAATGGAAGGTGAGCTCTACAGCTGGCAGCGCGGTGGACAGCCCAACGCTATCCGTA 488
QY 461 GCAACATAGCTATAGCAATGGCTCGCGTGGAGTAATGATTGAAGAAATGTATTATA 520
DB 489 GCAAGTGGGCATATCCATGGCTTGGCTGGATGTCAAGGCGCAAGAGTTCTACTTCA 548
QY 521 TTGATTCGGGGAAGAGTAGACGATGACGATGATGATGATGCTTCTACATATTCATCA 580
DB 549 TCGACACCAACACACGAGGATATGGCTTATGACTACAACTAGACGACCGCGCGGTGA 608
QY 581 GCAATCAACGGCATTTATTTACTTTTGAATA-----GCATGAAGTGCCTGGATATCCAG 634
DB 609 GCAACCAAGGTCATCTTCGATCTCAGGAGAGATTCGGCCCGCAAGACCATGTTTCCCTG 668
QY 635 ATGGTCAACAAATTTGATGAGGAGGGTAATTTATGGGTGTCGGCTTTTCCAGACACCGAA 694
DB 669 ATGCGATGACCGTAGACACCGATGGCAATATCTACGTGGCCACCTTCAATGGTGGCACCG 728

QY 695 TTATTAAATACAGTACCACCAACCGAAGTGTCTTACTGTATACCGTAAATATACCAGATC 754
DB 729 TCTTCAAGGTCAATCAAGCACCCTGTAATAATCTCTCTGGAGA---TCAAATTCACCA 785
QY 755 CTCAGCTACCTCTCTAGCAATTTGGCGTCCGAATTTGGATGAATGATGATTAACATCTG 814
DB 786 CCCAATACCTCGGTGGCTTTGGAGGTCCCACTTGGATATTTGTATGTGACAAACCG 845
QY 815 C 815
DB 846 C 846
RESULT 7
AF326959 992 bp mRNA linear INV 08-MAY-2001
LOCUS Calliphora vicina anterior fat body protein (APP) mRNA, partial
DEFINITION cds.
ACCESSION AF326959
VERSION AF326959.1 GI:13991404
KEYWORDS Calliphora vicina.
SOURCE Calliphora vicina
ORGANISM Calliphora vicina
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Calliphoridae; Calliphora.
REFERENCE 1 (bases 1 to 992)
AUTHORS Hansen, I.A., Meyer, S.R., Berlinger, M.J. and Scheller, K.
TITLE Identification of new interactors of the hexamerin receptor in the
biowfly Calliphora vicina
JOURNAL Unpublished
AUTHORS Hansen, I.A., Meyer, S.R., Berlinger, M.J. and Scheller, K.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Zoologie 1, University of Wuerzburg,
Biocenter Hubland, Wuerzburg 97074, Germany
FEATURES
source
1..992
/organism="Calliphora vicina"
/db_xref="taxon:7373"
/tissue_type="fat body"
/dev_stage="larvae"
<1..992
/gene="APP"
<1..804
/gene="APP"
/note="found as strong interactor of the hexamerin
receptor in a yeast 2-hybrid screen"
/codon_start=1
/product="anterior fat body protein"
/protein_id="AAK51353.1"
/db_xref="GI:13991405"
/translation="GYNNENKVKAKIEGDLAGFAPVVEGTTDQFVCSGRVVIV
QWCVSTAKVLKTLFEVQDGRFNTGNRDNKCDPRGLFAGTMRVYDEEHRNG
ELKYENGGEVEIKSDVIGSNGLANNKTKFYIDTIDYEVKDYDDETKASNP
KVFNLRKTPKDHLPDGMTIDTGNLYVATFNHGSIIYKINPTTGGVLLLEIKPECKO
ITSAFGSPNFHILFTVTSRSGEPHAGTYYKVTGLCAKGYPMTKIQL"
BASE COUNT 293 a 189 c 201 g 309 t
ORIGIN
Query Match 9.8%; Score 91; DB 3; Length 992;
Best Local Similarity 50.5%; Pred. No. 4.2e-12;
Matches 338; Conservative 0; Mismatches 310; Indels 21; Gaps 4;
QY 156 TAAAGTAGATAAATCGTGTCTTTCATTATTCCTGCTGATCCCTGCCCTTTTGT 215
DB 42 TGAAGGTGAGGATTTAGCTGGTGTTCCTGTCGCCGTGCAAGGTACACCGATCAATTTGC 101
QY 216 AGTCAAGTTTGAACGTGAATAGCCATTCTTACATGGATCGCGTTAGTGTGCACCTAC 275
DB 102 TGTGCTTCTGCTCGTCTGTTGTCATTGTACAAATGGCATGCTGTTCTGAGACCGCTAA 161
QY 276 AAGCATAGAAGCTATTGTTAATGTGCGAACACACATTTAAAT-----AACAGACT 326

```

Db 162 AGTATTGAAGACTTTGTCGAAGTGCACACCGGTGATGAAGCTTCTACTGCGCAACCGTTT 221
Qy 327 CAATGATGGCAAGACAGATCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGC 386
Db 222 CAATGATGGCAAGGTGATCTCTGTCGTTGCTTGGCGGTACCAATGCGCTTATGTCGG 281
Qy 387 TGGTCTCCCGTAGACCGGTCACTGCGAGTTTATATCAATTTAGGCGCTGATAAAGGT 446
Db 282 CGATGAATTCGAGCATCGTTA---TGGTGAATTTATACAAGTACGAAATGGAGCGAAGT 338
Qy 447 AAAAATGCACAGACACATAGCTATAGCAAAATGGGCTCGCTGGAGTAATGATTTGAA 506
Db 339 TGAGGTAATCAATCGGATGCGGTATTTCCAAATGGTTGGCTTGGATGAGAAGACTAA 398
Qy 507 GAAATGTATTATATTGATTCGGGGAAGAGAGTAGACGAGTACGATTAATGATGCTTC 566
Db 399 GAAATCTTATTACATGATACCAACCGGATTAAGGTTAAGGAGTATGATTATGATTTGA 458
Qy 567 TACATTATCCATGACGAATCAACGGCCATTAATTTACTTTTGAAG-----CATGAAGT 620
Db 459 AACTGGCAAGGCCCTCCCAATCCAAAGTTGTATTCATCTAGCGAAGACCGCCCAAGGA 518
Qy 621 GCCTGGATATCCAGATGCTCAACAATGATGAGGAGGTAAATTTATGGTTGCCGTTT 680
Db 519 TCATTTGTTGCCGATGCATGACCATGATGATGAGGCAATCTTTATGTGGCCACTTT 578
Qy 681 CCAAGGACAGCGAATTAATAAATCAGTACCACACCGGAAGTGTACTTGGATACCGT 740
Db 579 CAATGGCCATTCATCTACAAAATTAATCCACCACTGGCCAAAGTTTGTGGGAATC-- 636
Qy 741 AAAAATACCAGATCCTCAGTCACTCTGTAGACATTTGGCGGTCCGGAATTTGGATGAAC 800
Db 637 -AAATCCCTGCAACAAATCACCTCGGCTGCTTTTGGTGGTCCCAATTTCCATATTT 695
Qy 801 GCATGTAAC 809
Db 696 GTTTGTAAC 704

```

```

RESULT 8
E09237 LOCUS 1216 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA insert encoding Regucalcin.
ACCESSION E09237
VERSION E09237.1 GI:22025863
KEYWORDS JP 1995123985-A/1.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1216)
Yamaguchi, M.
DNA FRAGMENT CODING FOR REGUCALCIN
Patent: JP 1995123985-A 1 16-MAY-1995;
YAMAGUCHI MASAYOSHI, DAI ICHI PURE CHEM CO LTD
OS Rattus sp. (Rat)
PN JP 1995123985-A/1
PD 16-MAY-1995
PF 09-NOV-1993 JP 1993279349
PI YAMAGUCHI MASAYOSHI
PC C12N15/09,C12N1/21,(C12N1/21,C12R1.19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1216
FT CDS 80..979
FT /product="Regucalcin".
FT Location/Qualifiers
1..1216

```

```

FEATURES
source

```

```

/organism="Rattus sp."
/db_xref="taxon:10118"
BASE COUNT 338 a 231 c 309 g 338 t
ORIGIN
Query Match 9.8%; Score 90.4; DB 6; Length 1216;
Best Local Similarity 50.2%; Pred. No. 5.8e-12;
Matches 307; Conservative 0; Mismatches 292; Indels 12; Gaps 3;
Qy 298 GTCGAACACACATATAAAATAACAGACTCAATGATGGCAAGCAGATGCCCTTGGCAAT 357
Db 356 GTGATCAAGATAAGAAAACAATCGATTCAATGATGGAAAGTGGATCCTGCTGGGAGA 415
Qy 358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCGTA---GGACCGGTCACTGGC 414
Db 416 TACTTTGCTGTACCATGGCTGAGGAACACCCCGCAGCTGTTCTGGAGCGCACCAAGG 475
Qy 415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAATAATGCACGAGACAATAGCTATA 474
Db 476 TCCTTGTACTCCCTTTTCCCTGATGACAGTGTGAAGAAATACTTTAACCAAGTATATC 535
Qy 475 GCAATGGGCTCGGCTGGAGTAATGATTGGAAGAAAATGATATATTGATTTCGGGAAA 534
Db 536 TCCAAATGGTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCTGTCC 595
Qy 535 AGAAGAGTAGACGAGTAGATGATGCTTCTACATTTATCCATCAGCAANTCAACGGCCA 594
Db 596 TACACTGTGGATGCTTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCCGAGGACT 655
Qy 595 TTATTTACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGTCAGATGTCACAAACAATTTGATGAG 654
Db 656 GTTTTACAAGATGGAAGAAAGATGAA-----CAATCCAGATGGAATGTGCTATGATGTT 709
Qy 655 GAGGTATTTATGGTTGCCGTTTCCCAAGGACAGCGAATATTATAAATCAGTACCCAA 714
Db 710 GAGGGAAGCTTTGGTGGCTGTGTACAAATGGGAAGAGTAATTCGCCTAGATCCTGAG 769
Qy 715 CAACCGGAAGTGTACTGGATACCGTAAAAATACAGATCCTCAGGTCACTCTGTAGACA 774
Db 770 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGGCTGTTGATAAACAACCTTCATGCTGC 826
Qy 775 TTTGGCGTCCGAAATTTGGATGAAGTCACTGATCACTGCTGCTGCTTACGTTGACGCAC 834
Db 827 TTTGGAGGAAGGATTACTCTGAATGATGACGTGACATGTGCCAGGATGGGATGAGCCGC 886
Qy 835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGACGCTCTACAGAGTAACAGGTTTAGCGCTC 894
Db 887 GAAGGTCTTTTGAGCGACCTGATGCTGTTAACATTTTCAAGATAACAGGTTCTTGGGCTC 946
Qy 895 AAAGTTTCGC 905
Db 947 AAAGCAATTGC 957
RESULT 9
AB036903 LOCUS 1155 bp mRNA linear INV 12-JUL-2000
DEFINITION Sarcophaga peregrina AFP mRNA for anterior fat body protein, complete cds.
ACCESSION AB036903
VERSION AB036903.1 GI:8980620
KEYWORDS AFP; anterior fat body protein.
SOURCE Sarcophaga peregrina larva fat body cDNA to mRNA.
ORGANISM Sarcophaga peregrina
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Sarcophagidae; Sarcophaga.
1 (sites)
Nakajima, Y. and Natori, S.
IDENTIFICATION AND CHARACTERIZATION OF AN ANTERIOR FAT BODY PROTEIN
IN AN INSECT
JOURNAL J. Biochem. 127 (5), 901-908 (2000)
MEDLINE 20250851

```


[illegible]


```

Db      369 GTGGATGAAGATAAGAAAAACAATCGATTCAATGGAAGGTGGATCTGCTGGGAGA 428
Qy      358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCGCTA---GGACGGGTCACTGGC 414
Db      429 TACTTTGCTGTACCATGGCTGAGGAACCGCCCGACGCTGTTGGAGCGGCACCAAGGG 488
Qy      415 AGTTTATATCATTTAGGGGCTGATAAAAGGTAAAAATGCACGAGACCAACATAGCTATA 474
Db      489 TCCTTGCTACTCCCTTTTCTCGATCACAGTGCAGATGGAAGAAATACTTTAACCAAGTGGATATC 548
Qy      475 GCAATGGGCTCGCTGGAGTAATGATTTCGAAGAAATGTTATATATTGATTGCGGGAAA 534
Db      549 TCCAATGGTTGGATTGTCCTGGACCAATAAATCTTCTACTACATTCAGACGCTGTC 608
Qy      535 AGAAGAGTAGACAGTAGCATATTATGATGCTTCTACATTATCCATCCAGCAATCAACGCCA 594
Db      609 TACACTGTGGATGCTTTGACTATGACCTGCCACACAGCAGATTTCCACCGCAGGACT 668
Qy      595 TTATTTACTTTTGAAGACATGAAGTGCCTGGATATCCAGATGTCATAAACAATTGATGAG 654
Db      669 GTTTACAAAGATGGAAGAAAGATGAA-----CAAAATCCAGATGGAATGTGCATTGATGT 722
Qy      655 GAGGTAATTTATGGGTTCGGTTTCCAAAGGACAGCGAATTATTAAAAATCAGTACCCTAA 714
Db      723 GAGGGGAAGCTTTGGGGTGGCTGTGTACAATGGAGGAAGATAATTCGCCCTAGATCCTGAG 782
Qy      715 CAACCGGAAGTGTACTGGATACCGTAAAAATACACAGATCCTCAGGTCACTCTGTAGCA 774
Db      783 ACAGGGAAA---AGACTGCAACTGTGAAGTGCCTGTTGATAAAACAATTCATGCTGC 839
Qy      775 TTGGCGGCTCCGAATTTGGATGAACATGCATGTAAACATCTGCTGCTTTCAGCTTGACGAC 834
Db      840 TTTGGAGGGAAGGATTAATCTCTGAATACGATGACATGTGCCAGGGATGGGATGAGCGCC 899
Qy      835 AGTTCTTTNGACAAAAGTTTGTAGTTAATGGGACGCTCTACAGATTAACAGGTTTAGGCGTC 894
Db      900 GAAGGCTTTTGGAGGACGCTGTGCTGGTAACATTTTCAAGATAACAGGCTCTTGGGGTC 959
Qy      895 AAAGGTTTCGC 905
Db      960 AAAGGAATTGC 970

RESULT 11
RNSMP30A
LOCUS
DEFINITION R.norvegicus SMP30 mRNA for senescence marker protein-30.
ACCESSION X69021.1 S48114
VERSION X69021.1 GI:57254
KEYWORDS senescence marker protein-30.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS Fujita,T., Shirasawa,T., Uchida,K. and Maruyama,N.
TITLE Isolation of cDNA clone encoding rat senescence marker protein-30
(SMP30) and its tissue distribution
JOURNAL Biochim. Biophys. Acta 1132 (3), 297-305 (1992)
MEDLINE 93041931
PUBMED 1420310
REFERENCE
AUTHORS Fujita,T.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1992) T. Fujita, Dept. o Molecular Biology, Tokyo
Metropolitan Inst. of Gerontology, 35-2, Sakaecho, Itabashi-ku,
Tokyo 173, JAPAN
FEATURES
Location/Qualifiers
1..1594
/organism="Rattus norvegicus"
/strain="Wistar/Sic"

```

```

/db_xref="taxon:10116"
/dev_stage="6 months"
88..1588
/gene="SMP30"
CDS
88..987
/gene="SMP30"
/codon_start=1
/product="senescence marker protein-30"
/protein_id="CAA48786.1"
/db_xref="GI:57255"
/db_xref="SWISS-PROT:Q03336"
/translation="MSSIKIECVLRNRYRGSPVWEAEASKLLFLVDIPSKTVCRWDS
ISNRQVRGVDAVSSVALRQSGGIYVATIGTKFCALNWDQSIFILAWDEKKNRF
NDKRVDPAGRYFAGTAEETAPVLERHQGSLYSLFDPVKYKYFDQVDSLNGLDWSL
DHKIFYIDSLSYTDAFDYDLPTGQISNRRTVYKMEKDIQIDGMCIDYEGKLAAC
YNGRVLRLDPETGKRLQTVKLPVDKTTSCFCGKDYSEMYVTCARDGMSAEGLLRQP
DAGNIFKITGLGVKNIPYSYAG"
polyA_signal 1148..1153
/gene="SMP30"
polyA_signal 1566..1571
/gene="SMP30"
polyA_site 1588
/gene="SMP30"
BASE COUNT 466 a 298 c 366 g 464 t
ORIGIN
Query Match 9.3%; Score 85.6; DB 10; Length 1594;
Best Local Similarity 49.8%; Pred. No. 9e-11;
Matches 304; Conservative 0; Mismatches 295; Indels 12; Gaps 3;
Qy 298 GTGAAACACACATTTAAAAATAACAGACTCAATGATGGCAAGCAGATCCCTTTGGCAAT 357
Db 364 GTGGATGAAGATAAGAAAAACAATCGATTCAATGATGGAAGGTGGATCTGCTGGGAGA 423
Qy 358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCGCTA---GGACCGGTCACTGGC 414
Db 424 TACTTTGCTGTACCATGTGGAGAAACGGCCCCAGCTGTCTGAGCGCGCACCAAGG 483
Qy 415 AGTTTATATCATTTAGGGGCTGTATAAAGGTAAAAATGCAACGAGACAACATAGCTATA 474
Db 484 TCCTTGTACTCCCTTTTTCCTGTGATCAGATGCTGAAGAATACTTTGACCAAGTGATATC 543
Qy 475 GCAATGGGCTCGCTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTGCGGGAAA 534
Db 544 TCCAATGGTTGGATTGCTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 603
Qy 535 AGAAGAGTAGACGAGTAGCATTTATGATGCTTCTACATTATCCATCAGCAATCAACGCCA 594
Db 604 TACACTGTGATGCTTGTGACTATGACCTGCCAACAGGACAGATTTCCACCGCAGAACT 663
Qy 595 TTATTTACTTTTGAAGACGATGAAGTGCCTGGATATCCAGATGCTCAAAACAATTGATGAG 654
Db 664 GTTTACAAGATGGAAGAAAGATGAA-----CAATCCAGATGGAATGTGCATTGATGT 717
Qy 655 GAGGTAATTTATGGTTGGCTTTTCCAAAGGACAGCGAATATTATAAATCAGTACCCTAA 714
Db 718 GAGGGGAAGCTTTGGTGGCTGTTTACAATGGAGGAAGATAATTCGCTAGATCCTGAG 777
Qy 715 CAACCGGAAGTGTGTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACTCTCTAGCA 774
Db 778 ACAGGAAA---AGACTGCAAACTGTGAAGTTGCTGTTGATAAAACAACATTCATGCTGC 834
Qy 775 TTTGGCGTCCGAATTTGGATGAACATCTGCATTAACATCTGCTGCTTTCAGTCTTGAGAC 834
Db 835 TTTGGAGGAAGGATTACTCTGAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 894
Qy 835 AGTTCTTTNGACAAAAGTTTGTAGTTAATGGGACGCTCTACAGAGTAACAGGTTTAGGCGTC 894
Db 895 GAAGGCTCTTTGAGGACGCGCTGCTGGTGGTAACATTTTCAAGATAACAGGCTCTTGGGGTC 954
Qy 895 AAAGGTTTCGC 905
Db 955 AAAGGAATTGC 965

```

```

655 GAGGTAATTTATGGTGGCGTCTTTCAAGGACAGCGAATTAATTAATAATCAGTACCCAA 714
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
729 GAGGGAAGCTTTGGTGGCGCTGTACAAATGAGGAAGAGTAATTCGCTAGATCCTCAG 788
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
715 CAACGGGAAGTGTACTGGATACCGTAAATAATACAGATCCTCAGGTCACTCTGTAGCA 774
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
789 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGGCTGTGTGATAAACAACATTCATGCTGC 845
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
775 TTGGGGGTCGCAATTTGGATGAATGCAATGTAAGCATCTGCTGGTCTTTCAGCTTCACGAC 834
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
846 TTGGGGAAGGATTAATCTCTGAAATGTACGTGACATGTCAGGATGGGATGAGCGCC 905
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
835 AGTCTTTTTCACAAAAGTTAGTTAATGGCGACGTCTACAGAGTAACAGGTTTAGCGGTC 894
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
906 GAAGGCTCTTTTGGGAGCGCTGATGCTGTAACATTTTCAAGATAACAGGTCTTGGGGTC 965
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
895 AAAGGTTTCGC 905
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
966 AAAGGAATTGC 976
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13.
D86217
LOCUS
DEFINITION
Mus musculus mRNA for regucalcin, complete cds.
ACCESSION
D86217
VERSION
D86217.1 GI:1483151
KEYWORDS
calcium-binding protein; regucalcin.
SOURCE
Mus musculus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
1 (sites)
Murata,T. and Yamaguchi,M.
Molecular cloning of the cDNA coding for regucalcin and its mRNA
expression in mouse liver: the expression is stimulated by calcium
administration
Mol. Cell. Biochem. 173 (1-2), 127-133 (1997)
JOURNAL
MEDLINE
97422495
REFERENCE
2 (bases 1 to 1109)
Yamaguchi,M.
Direct Submission
Submitted (24-JUN-1996) Masayoshi Yamaguchi, University of
Shizuoka, Laboratory of Endocrinology and Molecular Metabolism,
Graduate School of Nutritional Sciences; 52-1 Yada, Shizuoka,
Shizuoka 422, Japan (Tel:054-264-5580, Fax:054-264-5580)
Sequence updated (15-Jul-1996) by: Masayoshi Yamaguchi.
COMMENT
Location/Qualifiers
FEATURES
source
1..1109
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="liver"
25..924
/codon_start=1
/product="regucalcin"
/protein_id="BAAL3046.1"
/db_xref="GI:1483152"
/translation="MSSIKVECLRENYRGESPVEEASQSLLFVDIPSKILCRWDT
VSNQVRVADPVSVALRQLGGYVATIGTKFCAALNWEQSVFILAMVEDKKNRRF
NDGKVDPAGRVFAETPAVLEHQSLSLFPDHSVKYFDQVDLSNGLDWSL
DHKIFYIDLSYTDVAFDYLQGOISNRRIVYKMEKDEQIPDGMCIDAEGLWAC
YNGRVRILDPETGKRLQTVKLVPLVDKTTSCCGGKDYSEMYVTTCARDGMAEGLLRP
DAGNIFKITGLGVGIAPYSYAG"

BASE COUNT 305 a 207 c 282 g 315 t
ORIGIN
Query Match 8.7%; Score 80.8; DB 10; Length 1109;
Best Local Similarity 49.3%; Pred. No. 1.5e-09;
Matches 301; Conservative 0; Mismatches 296; Indels 12; Gaps 3;

QY 298 GTCGAACACACATTAATAAATACAGACTCAATGATGCAAGCAGATCCCTTGGCAAT 357
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 GTGGATGAAGTAAGAAAACAAATCGATCAATGGAAGGTGATCCTGCTGGGAGA 434
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 CTATGACAGGTACAAATGGCTATTGACGGTGTCTCCCGGTA---GGACCGGTCACTGGC 414
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 435 TACTTTGCTGGTACCATTGCTGAGGAAACGGCCCGGCTGTCTGGAGCGGCACCAAGG 494
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 AGTTATATCATTTAGGGGCTATATAAAGGTAAAAATGCAGAGCAACATAGCTATA 474
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 495 TCCTTGATCCCTTTTCCCTGATCACAGTGTGAGAAATACTTTGACCAAGTGGATATC 554
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 475 GCAAAATGGCGTCGCTGGAGTAATGATTGGAAGAAATGTATATATTGATTCGGGGAAA 534
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 555 TCCAATGGTTTGGATTGGTCCCTGGACCAATAAAATCTTCTACTACATTCAGACGCTGTCC 614
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 535 AGAAGAGTAGACAGTACCAATATGATGCTTCTACATATCCATCAGCAATCAACGGCA 594
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 615 TACACTGTGGATGCCCTTTTCACTATGACCTGCGCAACAGGACAGATTTCCAAACCGCAACT 674
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 TTATTTTACTTTTCAAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAACAACATTCATGAG 654
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 675 GTTTACAAGATGGAANAAGATGAA-----CAATCCCAAGATGGAATGTGATGTT 728
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 301 GTGGATGAGATAGAAAAAATAATCGATTCAATGATGGGAAGGTGGATCTCTGCTGGGAGA 360
QY 358 CTATGGCAGGTACAAATGGCTATTGACGCTGGTCTCCCGTA---GGACCGGTCACTGGC 414
Db 361 TACTTTGCTGTACCATGCTAGGAACAGCGCCAGCTGTTCTTGACCGCACCAAGG 420
QY 415 AGTTTATATCAATTTAGGGGCTGATAAAGGTAAATGCAAGAGCAACATAGCTATA 474
Db 421 TCCCTGTACTCCCTCTTCTCCGTGATCACAGTGTGAAGAAATCTTTGACCAAGTGATATC 480
QY 475 GCAATGGCTCGGTGGAGTAATGATTGGAAGAAATGTTATATATGATTTCGGGGAA 534
Db 481 TCCATGTTGGATTGGTCCCTGGACCAATAAATCTTCTACTACATTTGACAGCTGTCC 540
QY 535 AGAAGAGTAGCAGTAGTATGATGCTTCTACATTTATCCATTCACAGCAATCAACGGCCA 594
Db 541 TACACTGTGGATGCTTTTGACTATGACCTACAAACAGCAGAGATTTCACCGCAGAAAT 600
QY 595 TTATTTACTTTTGAACATGAGTGGCTGGATATCCAGATGGTCAACAATGATGAG 654
Db 601 GTTTACAGATGGAAAGAGATGAA-----CAAAATCCAGATGGAATGTGCAATTTGATGCT 654
QY 655 GAGGTATTTATGGTGTCCGTTTCCAAAGCAGACGGAATTTATAAATCAGTACCCAA 714
Db 655 GAGGAAGCTATGGTGGCTGTACAAATGGAAGAGTAATTCGGCTGGATCCTGAG 714
QY 715 CAACCGGAAGTGTACTGGATPACGTAAATAACAGATCCTCAGTCACTCTCTGATCA 774
Db 715 ACAGGAAA---AGACTGCAAACTGTGAAGTTGCTGTGATAAACAACITTCATGCTGC 771
QY 775 TTTGGCGTCCGAATTTGGATGACTGATGATACATCTGCTGCTTTCAGCTTTGACGAC 834
Db 772 TTTGGAGGAAGATTTACTCTGAATGATGTGACCTGTGCCAGGATGGGTGAATGCT 831
QY 835 AGTCTTTNGACAAAGTTTAGTTAAATGGGACGCTCTACAGATTAACAGGTTTAGGCGTC 894
Db 832 GAAGCCCTTTGAGGACCGCTGATGCTGTAACATTTTCAAGATAACAGGCTCGGAGTC 891
QY 895 AAAGGTTTCG 905
Db 892 AAAGGAATTC 902

RESULT 14
MMU28937
LOCUS
DEFINITION Mus musculus senescence marker protein-30 (SMP30) mRNA, complete cds.
ACCESSION U28937
VERSION U28937.1 GI:1143999
KEYWORDS
SOURCE
ORGANISM Mus musculus.

REFERENCE
AUTHORS
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1573)
Fujita, T., Shirasawa, T. and Maruyama, N.
Isolation and characterization of genomic mouse senescence marker protein-30 (SMP30) Biochim. Biophys. Acta 1308 (1), 49-57 (1996)
JOURNAL
MEDLINE
PUBMED 96328264
8765750

REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1995) Takuji Shirasawa, Molecular Pathology, Tokyo Metropolitan Institute of Gerontology, 35-2, Sakaecho, Itabashi-ku, Tokyo 173, Japan
Location/Qualifiers
1. .1573
/organism="Mus musculus"
/strain="C57Black/6"
/db_xref="taxon:10090"

FEATURES
source

RESULT 15

/tissue_type="liver"
/dev_stage="adult"
1. .1573
/gene="SMP30"
73. .972
/gene="SMP30"
/codon_start=1
/product="senescence marker protein-30"
/protein_id="AAC52721.1"
/db_xref="GI:1144000"
/translation="MSSIKVECVLRNRYRCSESPVWEASQSLLEVDIPSKIICRWDT VSNQVRVADVPVSVVALRQLGVATIGTKFCALNWNQSVFLAWVDKKNRRE DKRVDPAGRYFACTMAETAPVLERHQSLYSLFPDHSYKVKFDQVDISGLDWSL DHKIFYIDSLSYTVDAFDYDQISNRIRIVYKMEKDEQIPDQMCIDAGKLVAC YNGRVRILDPETGKRLQTVKLPVDKTTSCCPGKDYSEMYVTCCARDGLNAEGLLRQPDAGNIFKITGLGVKGIAPYSYAG"
polya_signal 1132. .1137
/gene="SMP30"
1551. .1556
/gene="SMP30"
BASE COUNT 453 a 288 c 366 g 466 t
ORIGIN

Query Match 8.7%; Score 80.8; DB 10; Length 1573;
Best Local Similarity 49.3%; Pred. No. 1.4e-09;
Matches 301; Conservative 0; Mismatches 298; Indels 12; Gaps 3;
QY 298 CTCGAACACACATTTAAAAATAACAGCTCAATGATGCCAAGCAGATCCCTTCGCAAT 357
Db 349 GTGGATGAAGATAGAAAAAATAATCGATTCAATGATGGAGGTGGTCTGCTGGGAGA 408
QY 358 CTATGGCAGGTACATGGCTATTGACGCTGCTCTCCCGTA---GGACCGGTCACTGGC 414
Db 409 TACTTTGCTGGTACATGGCTGAGGAACAGCGCCAGCTGTTCTTGAGCGGCACCAAGG 468
QY 415 AGTTTATATCAATTTAGGGGCTGATAAAGGTAAATAATGCAGAGCAACATAGCTATA 474
Db 469 TCCTTGTACTCCCTCTTCTCCGTGATCACAGTGTGAAGAAATCTTTGACCAAGTGGATATC 528
QY 475 GCAATGGCTCGCGTGGAGTAATGATTGAAGAAATCTATTATATTGATTTCGGGAAA 534
Db 529 TCCAAATGGTTGGATTGGTCCCTGGACCATAAATCTCTTACTACATTGACAGCTGTCC 588
QY 535 AGAAGATGACGAGTAGCTATGATGCTTCTACATTTATCCATCAGCAATCAACGGCCA 594
Db 589 TACACTGTGGATGCTCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGATT 648
QY 595 TTATTTACTTTTGAAGACATGAAGTGGCTGGATATCCAGATGGTCAACAATTTGATGAG 654
Db 649 GTTTACAAGATGGAAAAAGATGAA-----CAAAATCCAGATGGAAATGTGCAATTTGATGCT 702
QY 655 GAGGTAATTTATGGTTGCGCTTTTCCAAAGCAGACGGAATTTATAAATCAGTACCCAA 714
Db 703 GAGGAAGCTATGGTGGCTGCTTTACAAATGGGAAGAGTAATTTGGCTGGATCCTGAG 762
QY 715 CAACCGGAAGTGTACTGGATACCGTAAATAATACAGATCCCTCAGGTCACTCTGTAGCA 774
Db 763 ACAGGAAA---AGACTCAAACTCTGAAGTTGCTGTTGATAAACAACATTTCACTGCTGC 819
QY 775 TTTGGCGTCCGAATTTGGATGAACTGATGTAACATCTGCTGGTCTTCAGCTTTCAGCAG 834
Db 820 TTTGGAGGAAGATTTACTCTGAAATGATGTGACCTGTGCCAGGATGGGTGGAATGCT 879
QY 835 AGTCTTTNGACAAAGTTTAGTTAAATGGGACGCTCTACAGATTAACAGGTTTAGCGCTC 894
Db 880 GAAGCCCTTTGAGGACCGCTGATGCTGTACATTTTCAAGATAACAGGCTCGGAGTC 939
QY 895 AAAGGTTTCG 905
Db 940 AAAGGAATTC 950

BC012710 1598 bp mRNA linear ROD 07-AUG-2002
 LOCUS Mus musculus, Similar to regucalcin, clone MGC:14006 IMAGE:4210374,
 DEFINITION mRNA, complete cds.
 ACCESSION BC012710 GI:15215230
 VERSION BC012710.1
 KEYWORDS MGC.
 ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1598)
 Strausberg, R.
 Direct Submission
 Submitted (15-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 18 Row: n Column: 22
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6677738.

FEATURES

Location/Qualifiers
 1..1598
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:14006 IMAGE:4210374"
 /tissue_type="Liver, normal. 5 month old male mouse."
 /clone_lib="NCI_CGAP_Li9"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 86..985
 /codon_start=1
 /product="Similar to regucalcin"
 /protein_id="AAH12710.1"
 /db_xref="GI:15215231"
 /translation="MSSIKVECLRENYRCGSPVWEEASQSLFLVDIPSKIICRWDT
 VSNQVRVADAPVSSVALRQLGGVATIGTKFCALNWNQSVFLAMVDKNNRFL
 NDKVDPAGRYFAGTMAETAPVLERHQSLSLFPDHSVKYFDQVDISNGLDWSL
 DHKIFYIDSLSYTDADYDLOTQISNRRIVYKMKDEQIPDGMCI DAEGKLWVAC
 YNGRVRILDPETGRKRLQVLPVDTKTCSCFPGGDYSEMIYVICARDGLNAGLLRQP
 DAGNFKITGLGVKGIAPYSYAG"
 BASE COUNT 466 a 293 c 370 g 469 t
 ORIGIN

Query Match 8.7%; Score 80.8; DB 10; Length 1598;
 Best Local Similarity 49.3%; Pred. No. 1.4e-09;
 Matches 301; Conservative 0; Mismatches 298; Indels 12; Gaps 3;
 QY 298 GTGAACACACATTAAAAATACAGACTCAATGATGGCAAGACAGATCCCTTGGCAAT 357
 DB 362 GTGGATGAAGATAAGAAAAAATATCGATTCAATGATGGGAAGGTGCTGCTGGAGA 421
 QY 358 CTATGACAGGTACATGGCTATTGACGTGGTCTCCCGTA---GGACCGGTCACTGGC 414
 DB 422 TACTTTGCTGGTACCATGGCTAGGAACGGCCCGCCAGCTGTCTTTGAGCGGCCCAAGGG 481
 QY 415 AGTTTATATCTTTAGGGCTGATAAAAAAGGTAATAAGTGAACAGAGACATAGCTATA 474

Db 482 TCCTTCTACTCCCTCTTTCTCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 541
 QY 475 GCAAAATGGGCTCCGCTGGAGTATGATTTGAGAAAATGTATTATATTGATTTCGGGAAA 534
 Db 542 TCCAATGGTTGGATTGGTCCCTGGACCATAAATCTTCTACTACATTGACAGCCTGTCC 601
 QY 535 AGAAGAGTACAGCAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGCCA 594
 Db 602 TACACTGTGGATGCCCTTTGACTATGACCTACAACACAGGACAGATTTCCAACCCAGAAT 661
 QY 595 TTATTTACTTTTGAAGAAGCAATGAAGTCCCTGGATATCCAGATGGTCAAAACAATTGATGAG 654
 Db 662 GTTTACAAGATGCAAAAAGATGAA-----CAAAATCCAGATGGAATGTGCAATTGATGCT 715
 QY 655 GAGGTAATTTATGGTGGTGGCTTTTCCAAAGGACAGCGAATTTATAAATCAGTACCCAA 714
 Db 716 GAGGGAAGCTATGGGTGGCTGTACATGGAGGAAGAGTAATTCGCCTGGATCCTGAG 775
 QY 715 CAACCGGAAGTGTACTGGATACCGTAAAAAATACCAGATCCTCAGGTCACTCTGTAGCA 774
 Db 776 ACAGGGAAA--AGACTGCAAACTGTGAAGTTGCTTGTGATAAAACAACCTTCATGCTGC 832
 QY 775 TTTGGCGGTCGGAATTTGGGATGAACATCTCTGCTGTCTTCAAGTTTGAGGAC 834
 Db 833 TTTGGAGGGAAGATTAATCTCTGAATGTATGTGACCTGTCCAGGGATGGTTGAATGCT 892
 QY 835 AGTTCTTTNGACAAAAGTTTATAGTTAATGGGCACCTCTACAGAGTAACAGGTTTAGCGCTC 894
 Db 893 GAAGGCCCTTTTGAGGCAGCCTGATGCTGGTAAACATTTTCAAGATAACAGGCTCTCGGAGTC 952
 QY 895 AAAGGTTTCGC 905
 Db 953 AAAGGAATTGC 963

Search completed: August 1, 2003, 12:51:16
 Job time : 2658 secs